ZKFault: Fault Attack Analysis on Zero-Knowledge Based Post-Quantum Digital Signature Schemes

Puja Mondal, Supriya Adhikary, Suparna Kundu, and Angshuman Karmakar

Asiacrypt 2024, Kolkata, India

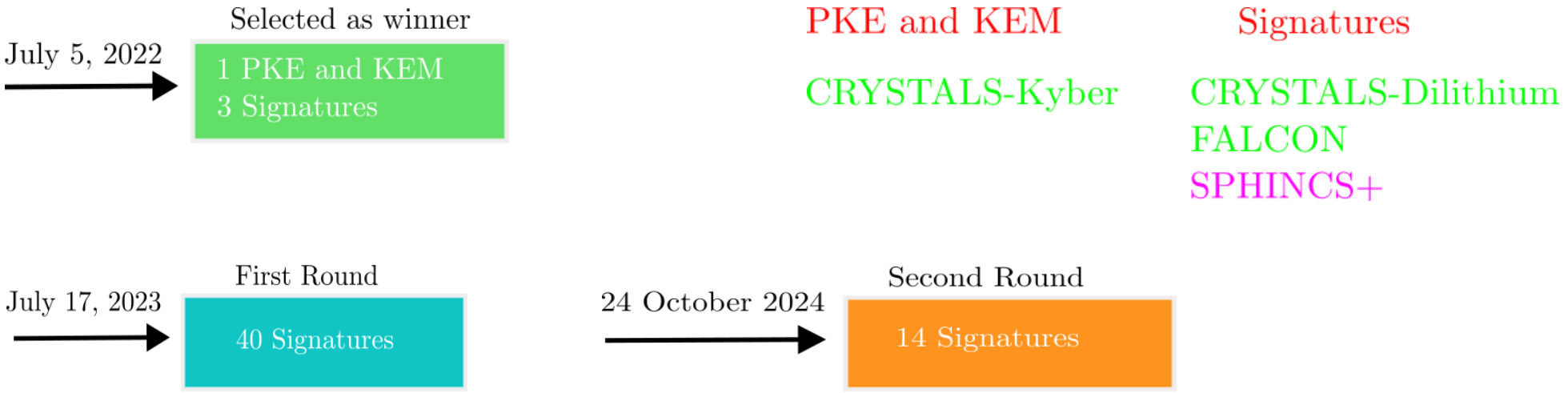
NIST's Additional Digital Signature

PKE and KEM	Signatures
CRYSTALS-Kyber	CRYSTALS-Dilithium
	FALCON
	SPHINCS+

NIST's Additional Digital Signature

PKE and KEM	Signatures
CRYSTALS-Kyber	CRYSTALS-Dilithium
	FALCON
	SPHINCS+

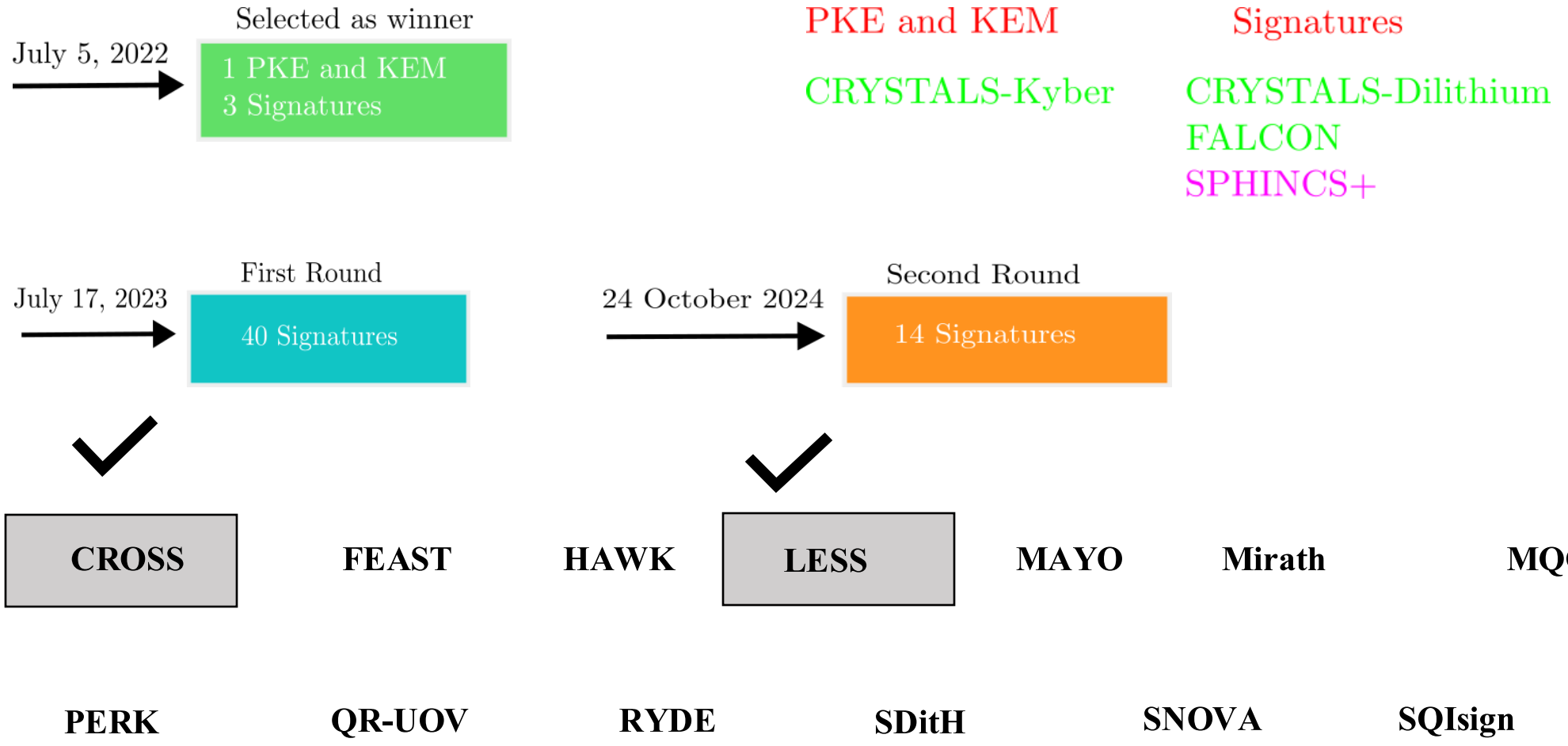
NIST's Additional Digital Signature



CROSS **FEAST** **HAWK** **LESS** **MAYO** **Mirath** **MQOM**

PERK **QR-UOV** **RYDE** **SDitH** **SNOVA** **SQIsign** **UOV**

NIST's Additional Digital Signature



Interactive-Zero-Knowledge (ZK) Framework

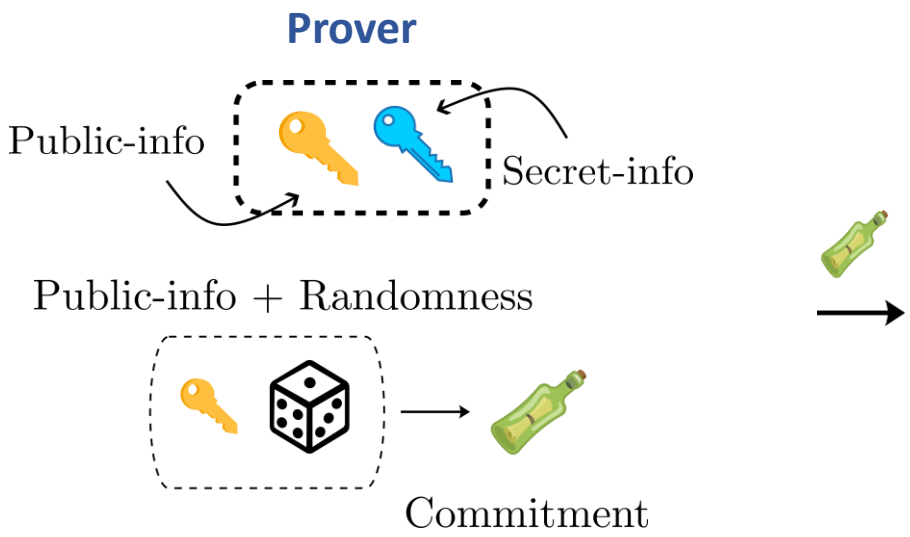
Prover **Verifier**

Interactive-Zero-Knowledge (ZK) Framework

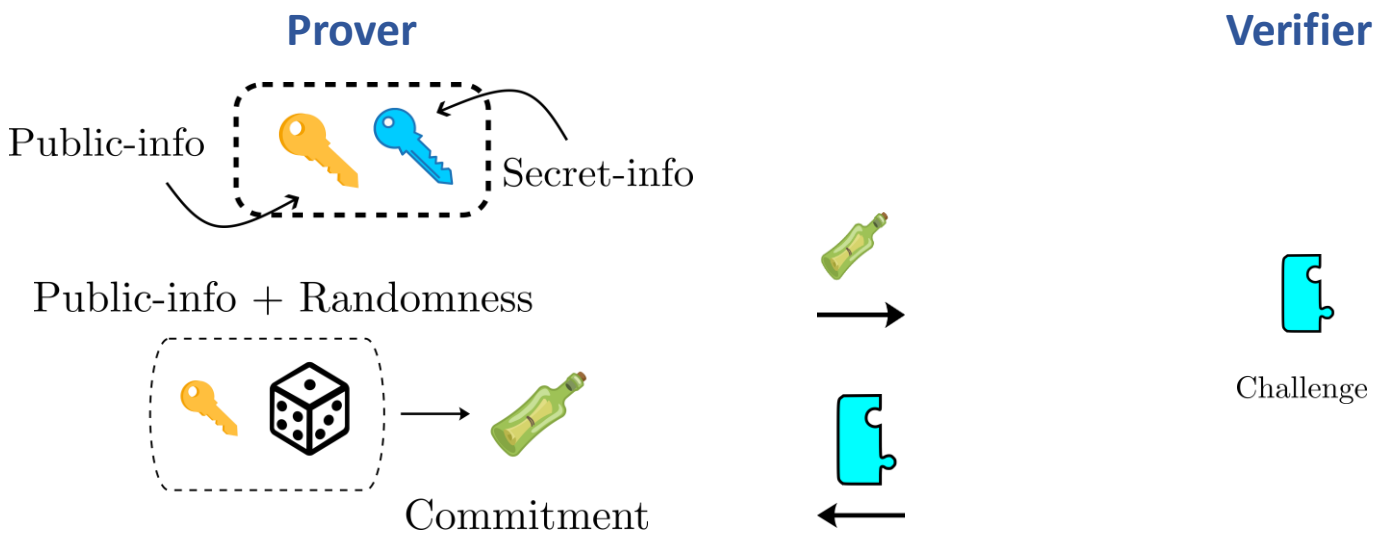
Prover

Verifier

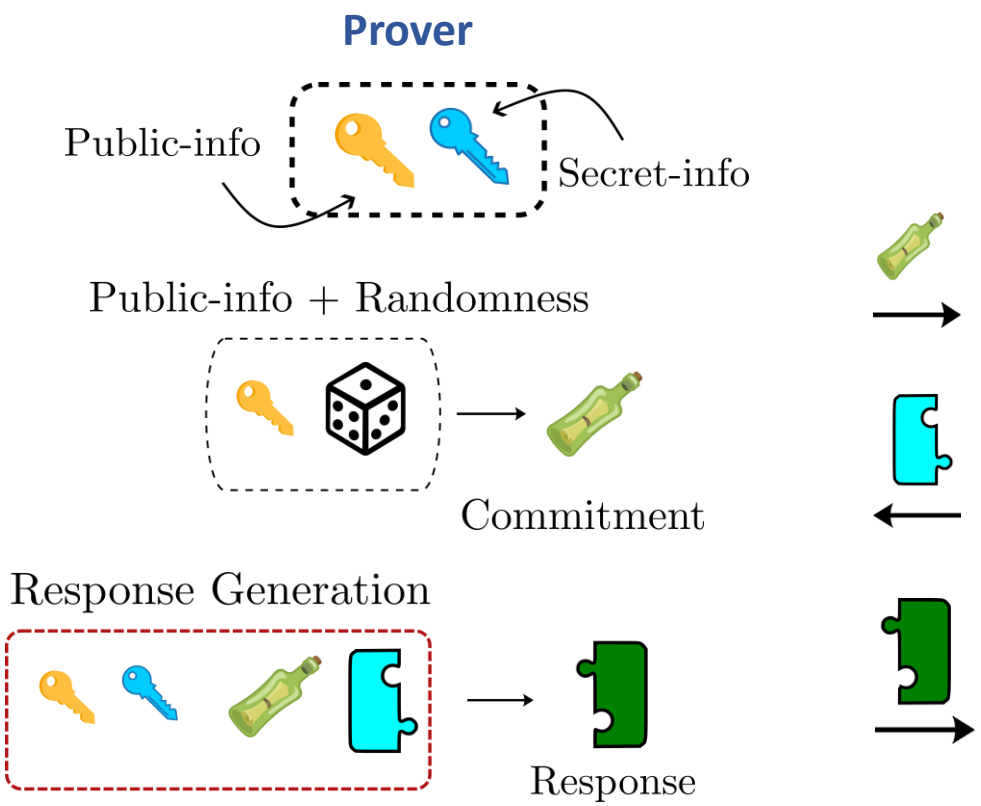
Interactive-Zero-Knowledge (ZK) Framework



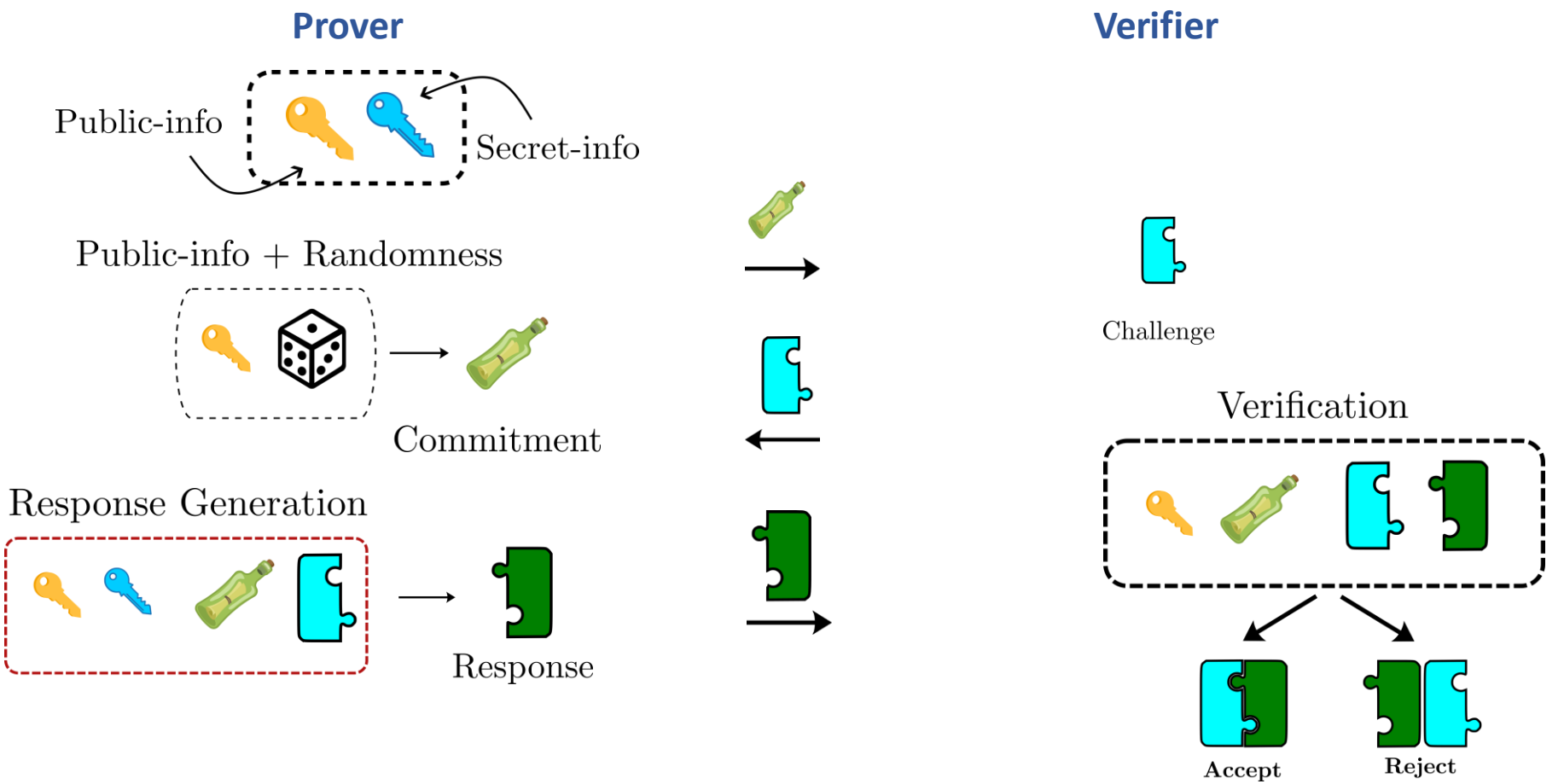
Interactive-Zero-Knowledge (ZK) Framework



Interactive-Zero-Knowledge (ZK) Framework

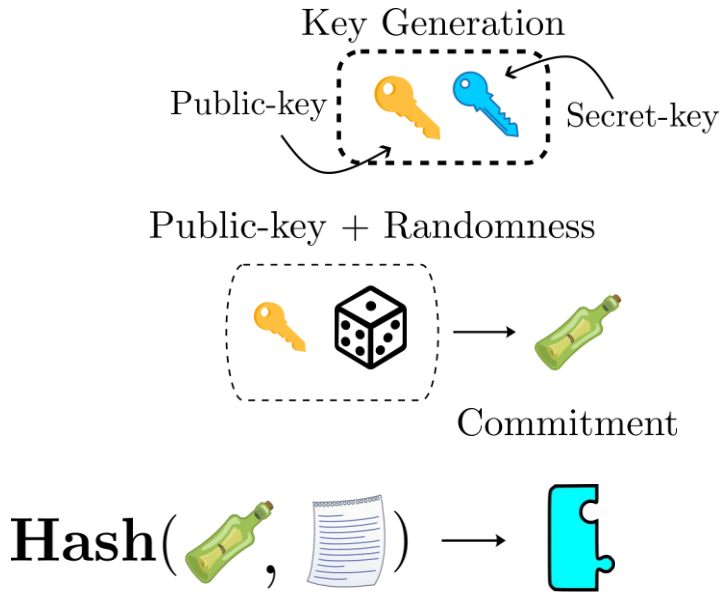


Interactive-Zero-Knowledge (ZK) Framework



ZK with Fiat-Shamir Transformation

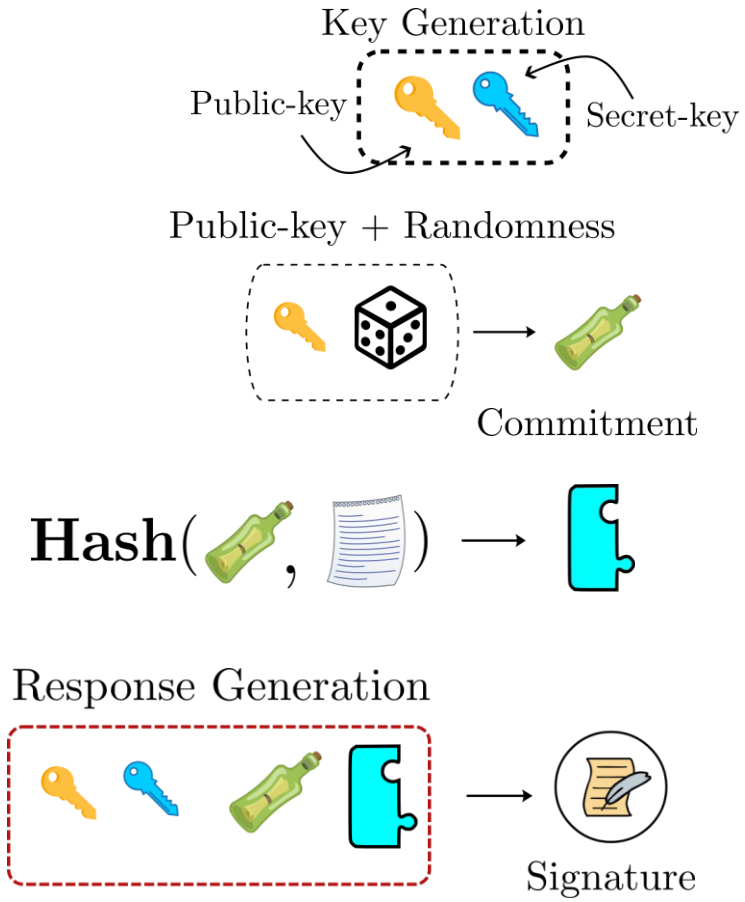
Signer



Verifier

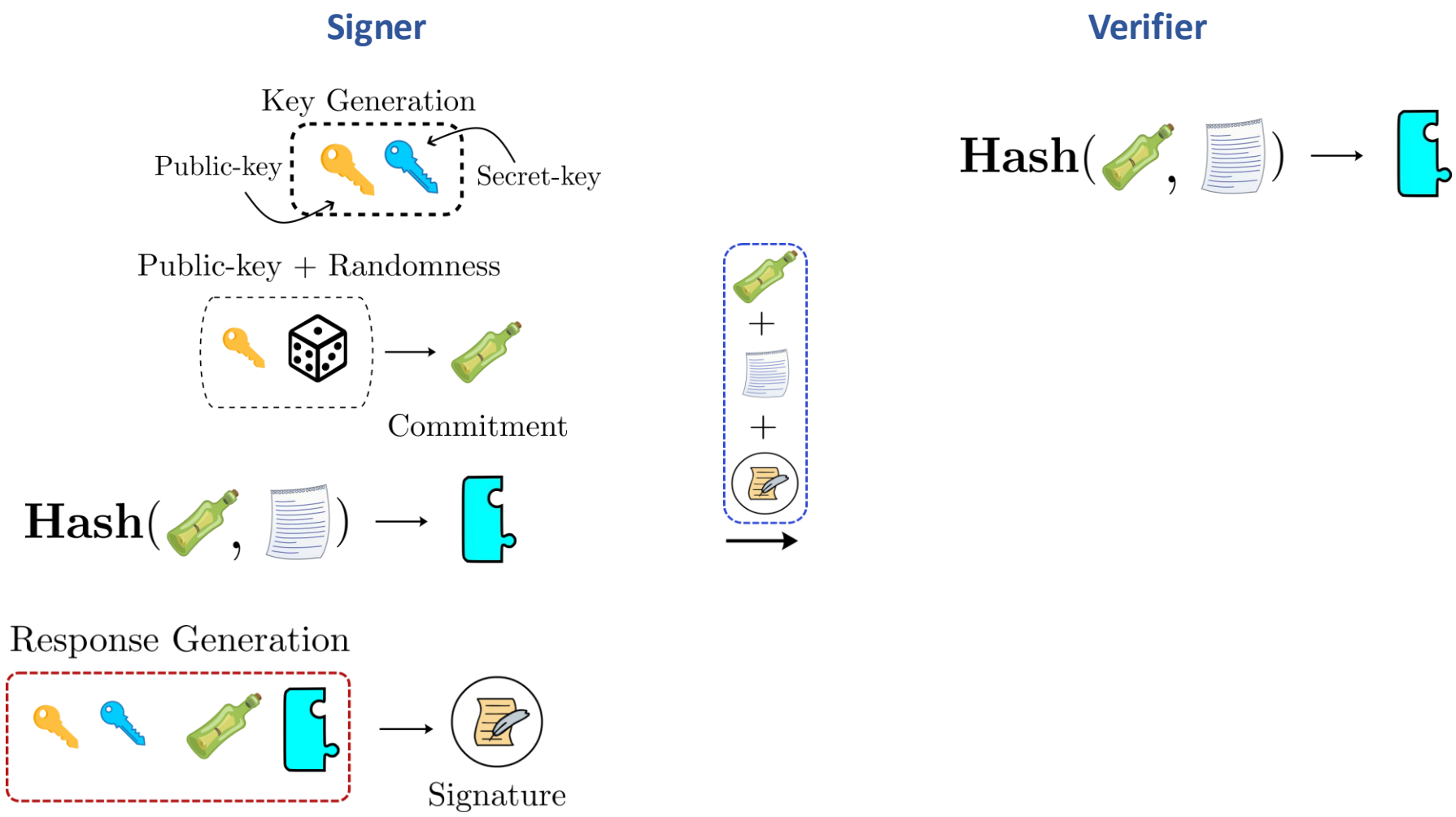
ZK with Fiat-Shamir Transformation

Signer

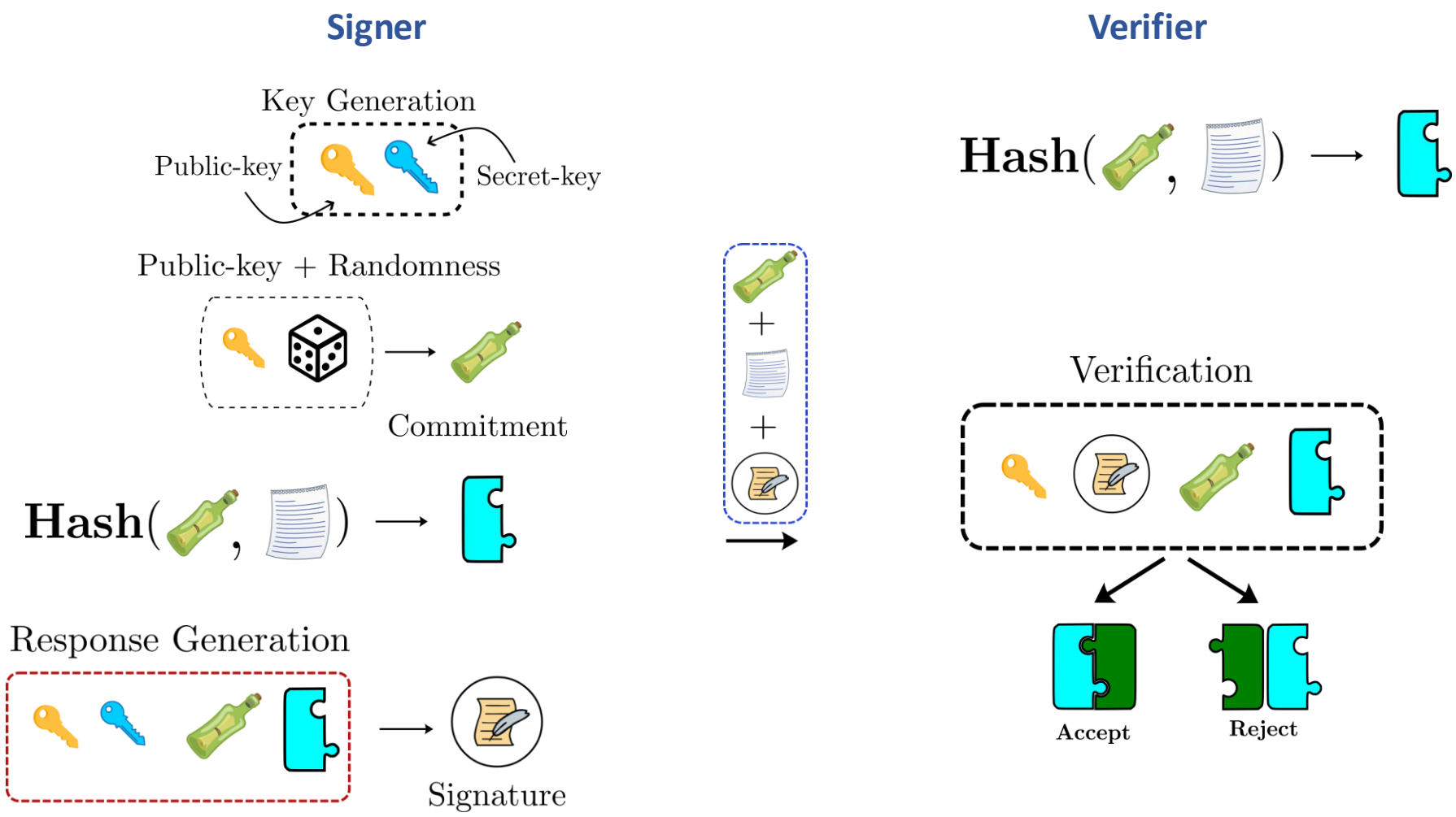


Verifier

ZK with Fiat-Shamir Transformation



ZK with Fiat-Shamir Transformation



LESS

Signer **Verifier**

🔧 : $\boldsymbol{Q}_1, \boldsymbol{Q}_2, \cdots, \boldsymbol{Q}_s$

🔑 : $\boldsymbol{G}_0, \boldsymbol{G}_1 = S(\boldsymbol{G}_0\boldsymbol{Q}_1), \boldsymbol{G}_2, \cdots, \boldsymbol{G}_s$

LESS

Signer **Verifier**

🔧 : $\boldsymbol{Q}_1, \boldsymbol{Q}_2, \cdots, \boldsymbol{Q}_s$

🔑 : $\boldsymbol{G}_0, \boldsymbol{G}_1 = S(\boldsymbol{G}_0\boldsymbol{Q}_1), \boldsymbol{G}_2, \cdots, \boldsymbol{G}_s$

$\textbf{Sample} : \boldsymbol{Q}'_1, \boldsymbol{Q}'_2, \cdots, \boldsymbol{Q}'_t$

$\boldsymbol{G}'_1 = S(\boldsymbol{G}_0\boldsymbol{Q}'_1), \boldsymbol{G}'_2, \cdots, \boldsymbol{G}'_t$

LESS

Signer **Verifier**

🔑 : $\boldsymbol{Q}_1, \boldsymbol{Q}_2, \cdots, \boldsymbol{Q}_s$

🔑 : $\boldsymbol{G}_0, \boldsymbol{G}_1 = S(\boldsymbol{G}_0\boldsymbol{Q}_1), \boldsymbol{G}_2, \cdots, \boldsymbol{G}_s$

$\mathbf{Sample}: \boldsymbol{Q}'_1, \boldsymbol{Q}'_2, \cdots, \boldsymbol{Q}'_t$

$\boldsymbol{G}'_1 = S(\boldsymbol{G}_0\boldsymbol{Q}'_1), \boldsymbol{G}'_2, \cdots, \boldsymbol{G}'_t$

🍾 : $\mathbf{Hash}\left(\boldsymbol{G}'_1, \boldsymbol{G}'_2, \cdots, \boldsymbol{G}'_t\right)$

Commitment

Challenge

$\mathbf{Hash}($ 🍾 , 📄 $) \longrightarrow$ 🧩 : $\mathbf{c_1}, \mathbf{c_2}, \cdots, \mathbf{c_t} \in \{\mathbf{0}, \cdots, \mathbf{s}\}$

LESS

Signer **Verifier**

$: Q_1, Q_2, \cdots, Q_s$

$: G_0, G_1 = S(G_0 Q_1), G_2, \cdots, G_s$

$\textbf{Sample} : Q'_1, Q'_2, \cdots, Q'_t$

$G'_1 = S(G_0 Q'_1), G'_2, \cdots, G'_t$

$: \textbf{Hash}(G'_1, G'_2, \cdots, G'_t)$

Commitment

Challenge

$\textbf{Hash}(\quad , \quad) \longrightarrow \quad : \mathbf{c_1}, \mathbf{c_2}, \cdots, \mathbf{c_t} \in \{\mathbf{0}, \cdots, \mathbf{s}\}$

$\mathbf{c}_i \neq \mathbf{0}$: $Q_{\mathbf{c}_i}^{-1} Q'_i$

$: R_i$

$\mathbf{c}_i = \mathbf{0}$: Q'_i

LESS

Signer

$$\text{🔑} : Q_1, Q_2, \cdots, Q_s$$

$$\text{🔑} : G_0, G_1 = S(G_0 Q_1), G_2, \cdots, G_s$$

$$\text{Sample} : Q'_1, Q'_2, \cdots, Q'_t$$
$$G'_1 = S(G_0 Q'_1), G'_2, \cdots, G'_t$$

$$\text{🍾} : \mathbf{Hash}(G'_1, G'_2, \cdots, G'_t)$$

Commitment

Challenge

$$\mathbf{Hash}(\text{🍾}, \text{📄}) \longrightarrow \text{🧩} : \mathbf{c_1}, \mathbf{c_2}, \cdots, \mathbf{c_t} \in \{\mathbf{0}, \cdots, \mathbf{s}\}$$

$$\mathbf{c}_i \neq \mathbf{0} : Q_{\mathbf{c}_i}^{-1} Q'_i$$

$$\text{✒} : R_i$$

$$\mathbf{c}_i = \mathbf{0} : Q'_i$$

Verifier

Challenge

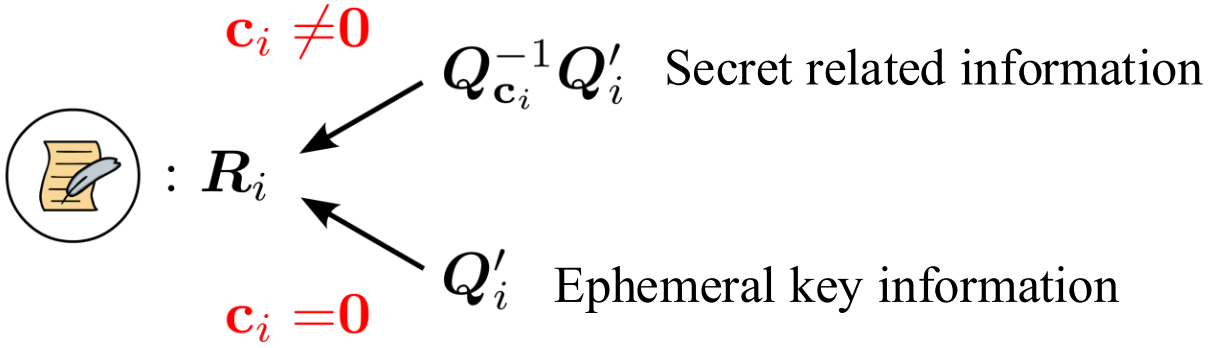
$$\mathbf{Hash}(\text{🍾}, \text{📄}) \longrightarrow \text{🧩}$$

$$G'_i = S(G_{\mathbf{c}_i} R_i)$$

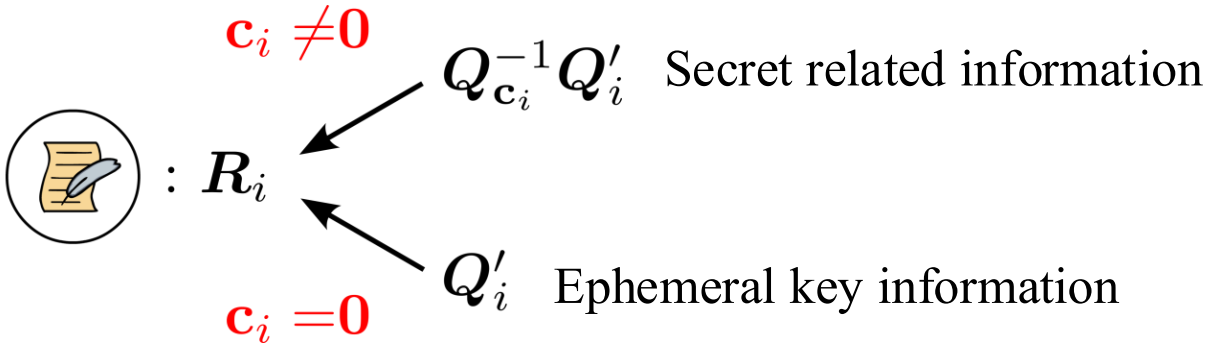
$$\mathbf{Check}:$$

$$\text{🍾} = \mathbf{Hash}(G'_1, G'_2, \cdots, G'_t)$$

Target of Our Fault Attack

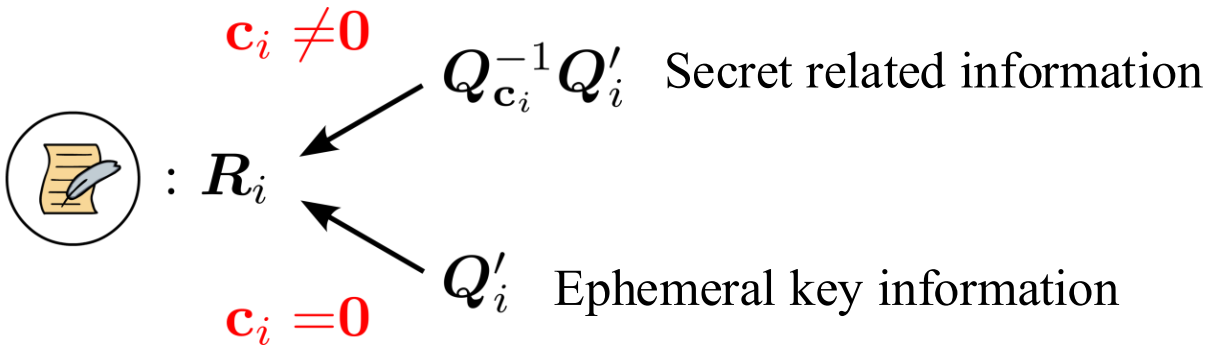


Target of Our Fault Attack



- For a fixed $\mathbf{c_i}$, we will not get both $\mathbf{Q'_i}$ or $\mathbf{Q^{-1}_{ci}Q'_i}$

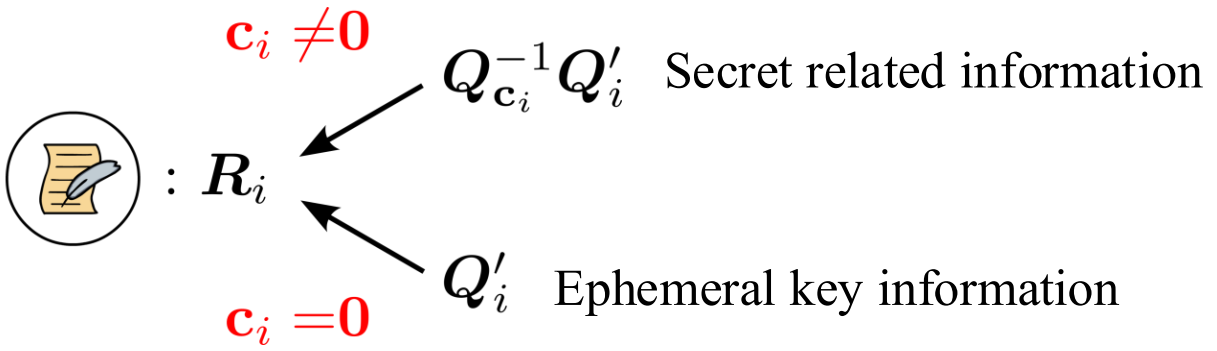
Target of Our Fault Attack



- For a fixed $\mathbf{c_i}$, we will not get both $\mathbf{Q'_i}$ or $\mathbf{Q^{-1}_{ci}Q'_i}$

- If we have both $\mathbf{Q'_i}$ and $\mathbf{Q^{-1}_{ci}Q'_i}$, then the secret $\mathbf{Q_{ci}=Q'_i(Q^{-1}_{ci}Q'_i)^{-1}}$ will be recovered.

Target of Our Fault Attack



- For a fixed $\mathbf{c_i}$, we will not get both $\mathbf{Q'_i}$ or $\mathbf{Q^{-1}_{ci}Q'_i}$

- If we have both $\mathbf{Q'_i}$ and $\mathbf{Q^{-1}_{ci}Q'_i}$, then the secret $\mathbf{Q_{ci}=Q'_i(Q^{-1}_{ci}Q'_i)^{-1}}$ will be recovered.

Target

Get a pair $(\mathbf{Q'_i, Q^{-1}_{ci}Q'_i})$

Fault Assumption and Result

- Our fault assumption is we can change the value of a location from 1 to 0

Fault Assumption and Result

- Our fault assumption is we can change the value of a location from 1 to 0

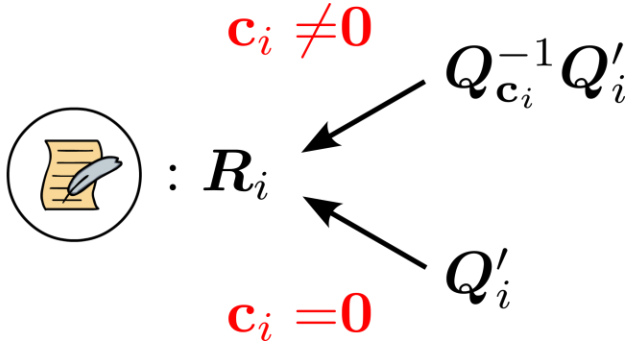
- This can be achieved by any one of
  - bit flip fault
  - stuck at zero fault
  - instruction skip fault

Fault Assumption and Result

- Our fault assumption is we can change the value of a location from 1 to 0
- This can be achieved by any one of
  - bit flip fault
  - stuck at zero fault
  - instruction skip fault

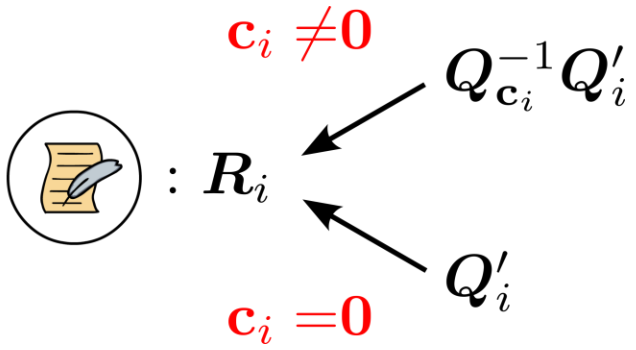
Scheme name	Security level	Parameter set	Required faults for complete secret recovery
LESS	1	Less-1b	**1**
		Less-1i	**1**
		Less-1s	**2**
	3	Less-3b	**1**
		Less-3s	**1**
	5	Less-5b	**1**
		Less-5s	**1**
CROSS	1, 3, 5	CROSS-R-SDP	**1**
		CROSS-R-SDPG	**1**

Overview of LESS



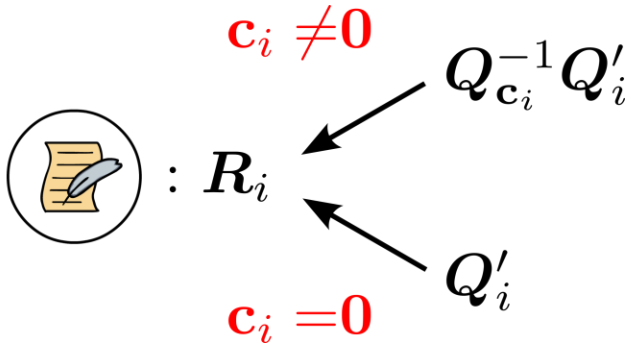
Overview of LESS

- Each $\mathbf{Q'_i}$ is generated from a random seed $\mathbf{seed_i}$



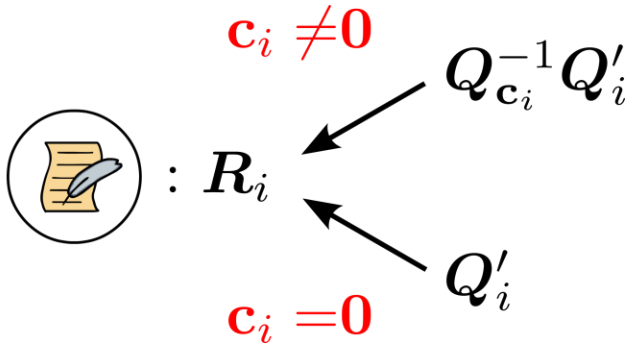
Overview of LESS

- Each $\mathbf{Q'_i}$ is generated from a random seed $\mathbf{seed_i}$
- The challenge $\mathbf{c=c_1,c_2,..,c_t}$ is of fixed weight w



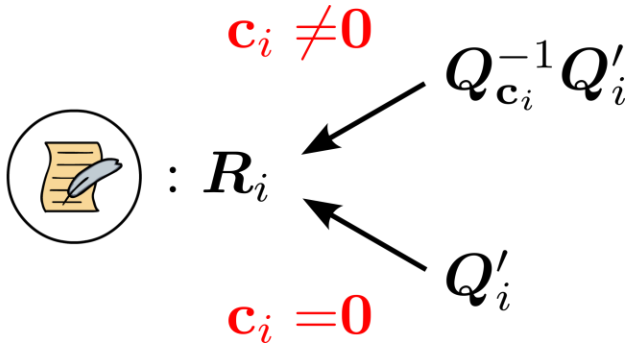
Overview of LESS

- Each $\mathbf{Q'_i}$ is generated from a random seed $\mathbf{seed_i}$
- The challenge $\mathbf{c=c_1,c_2,..,c_t}$ is of fixed weight w
- The response must contain information of (t-w) many seeds of $\mathbf{Q'_i}$



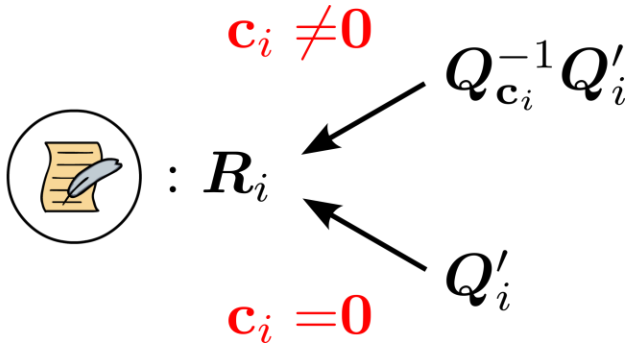
Overview of LESS

- Each $\mathbf{Q'_i}$ is generated from a random seed $\mathbf{seed_i}$
- The challenge $\mathbf{c=c_1,c_2,..,c_t}$ is of fixed weight w
- The response must contain information of (t-w) many seeds of $\mathbf{Q'_i}$
- For parameter, LESS-5b signature size 74960 Bytes



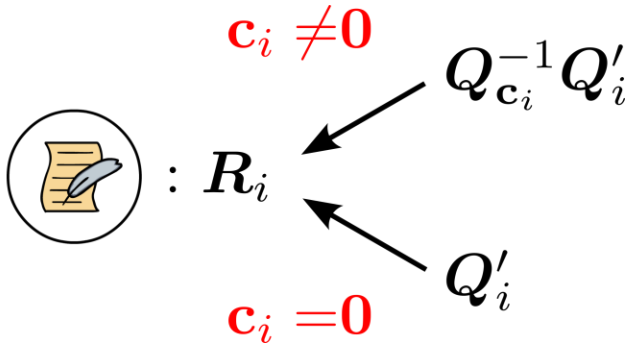
Overview of LESS

- Each $\mathbf{Q'_i}$ is generated from a random seed $\mathbf{seed_i}$
- The challenge $\mathbf{c=c_1,c_2,..,c_t}$ is of fixed weight w
- The response must contain information of (t-w) many seeds of $\mathbf{Q'_i}$
- For parameter, LESS-5b signature size 74960 Bytes
- LESS uses Goldreich-Goldwasser-Micali (GGM) tree construction



Overview of LESS

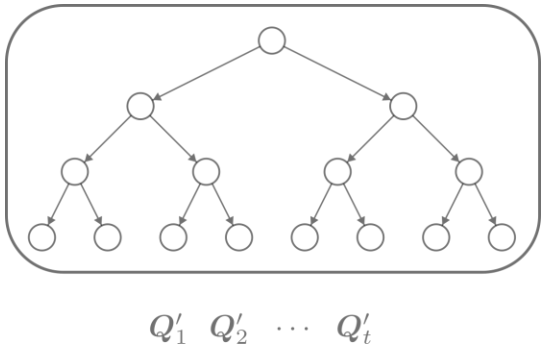
- Each $\mathbf{Q'_i}$ is generated from a random seed $\mathbf{seed_i}$
- The challenge $\mathbf{c=c_1,c_2,..,c_t}$ is of fixed weight w
- The response must contain information of (t-w) many seeds of $\mathbf{Q'_i}$
- For parameter, LESS-5b signature size 74960 Bytes
- LESS uses Goldreich-Goldwasser-Micali (GGM) tree construction
- Compressed signatures size: ~32500 Bytes (~57% reduced)



Summary of Signature Compression

Step 1

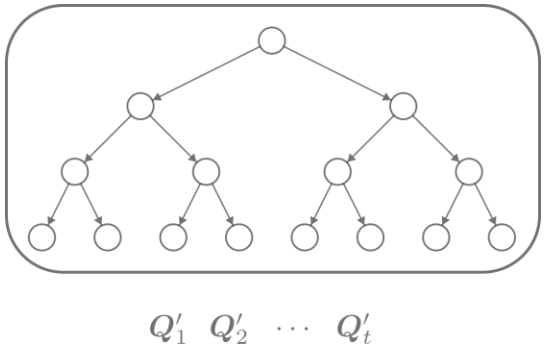
Generate a GGM tree (seedtree)
and generate all the ephimeral
keys Q_1', Q_2', … ,Q_t'



Summary of Signature Compression

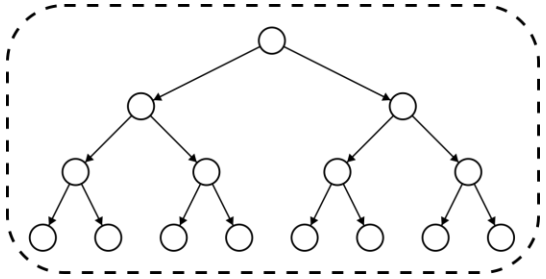
Step 1

Generate a GGM tree (seedtree) and generate all the ephimeral keys Q_1', Q_2', … ,Q_t'



Step 2

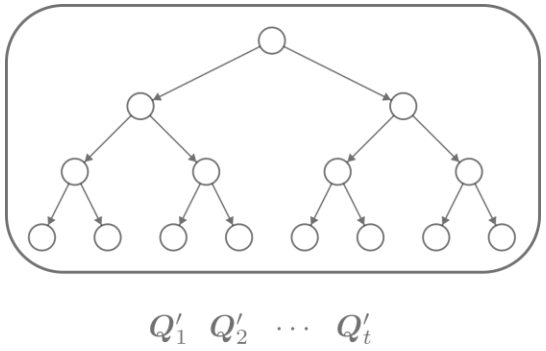
Generate a reference tree based on the challenge



Summary of Signature Compression

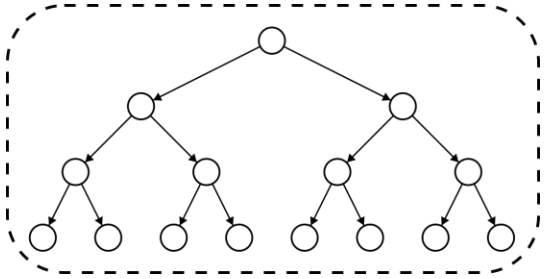
Step 1

Generate a GGM tree (seedtree) and generate all the ephimeral keys Q_1', Q_2', … ,Q_t'



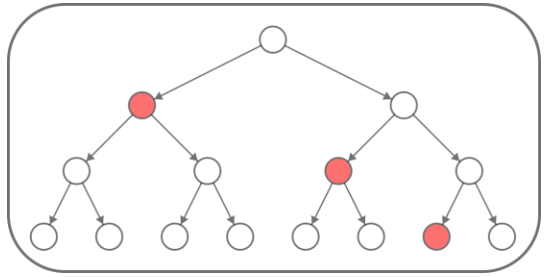
Step 2

Generate a reference tree based on the challenge



Step 3

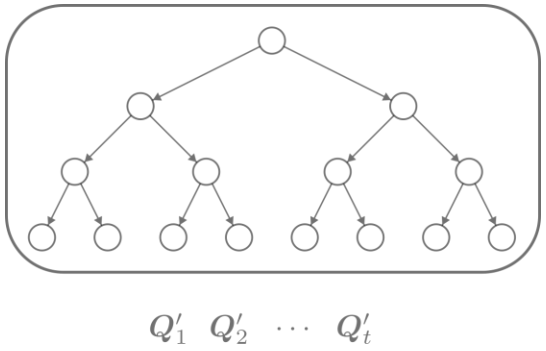
Decide the nodes of the seedtree to be published based on reference tree



Summary of Signature Compression

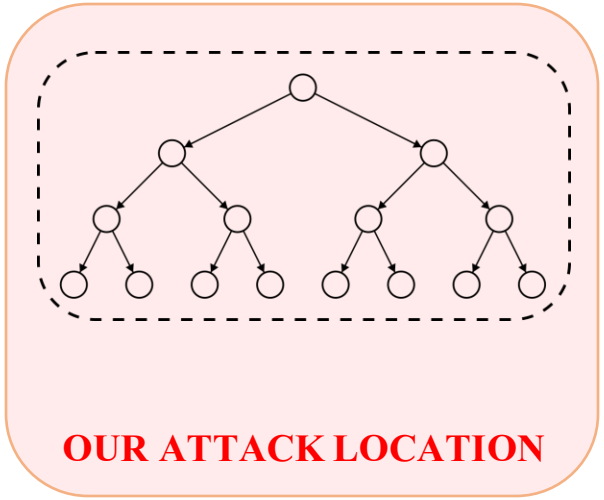
Step 1

Generate a GGM tree (seedtree) and generate all the ephimeral keys Q_1', Q_2', … ,Q_t'



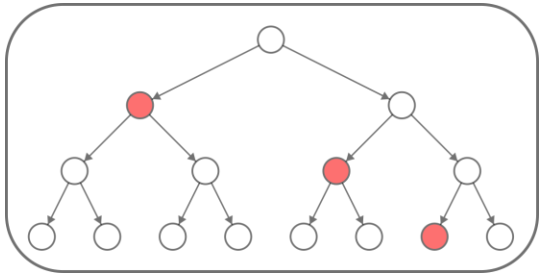
Step 2

Generate a reference tree based on the challenge

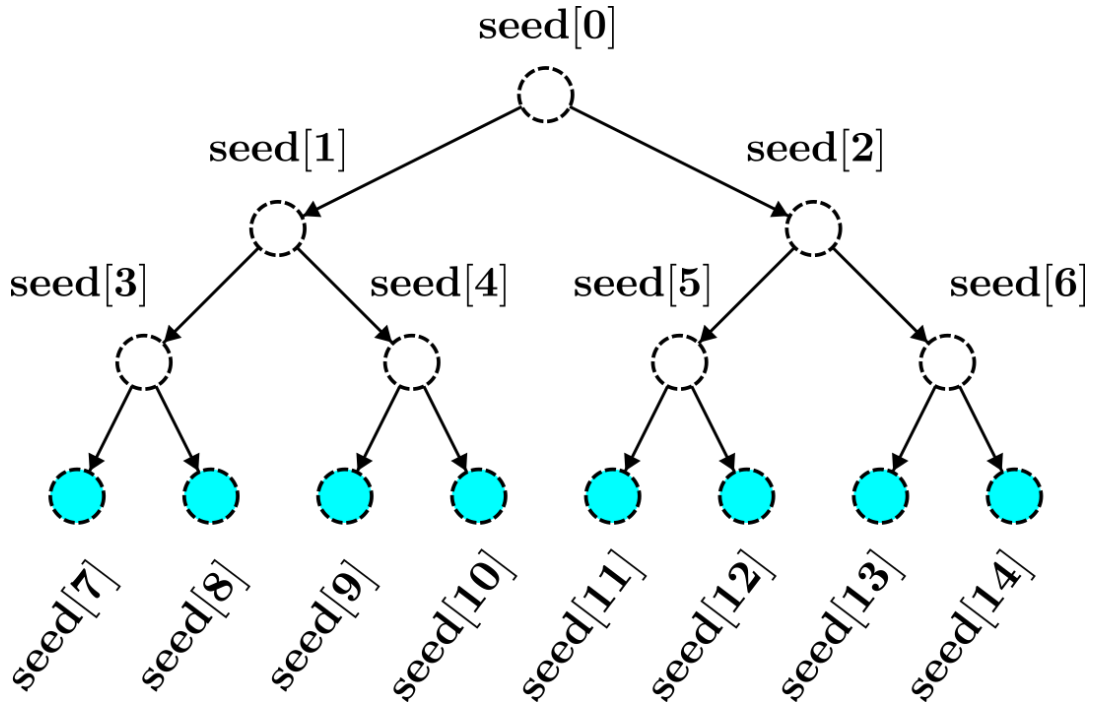


Step 3

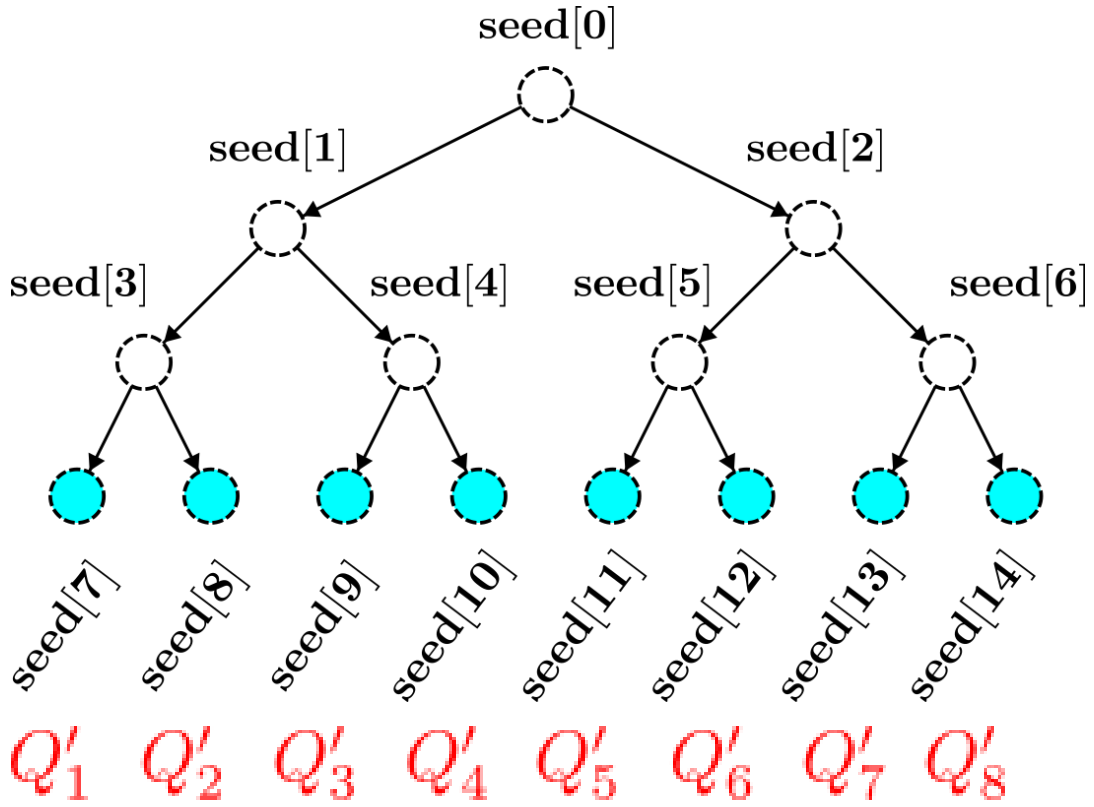
Decide the nodes of the seedtree to be published based on reference tree



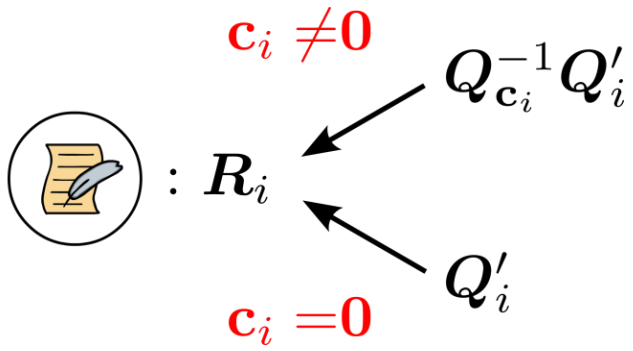
Seed Tree Generation



Seed Tree Generation

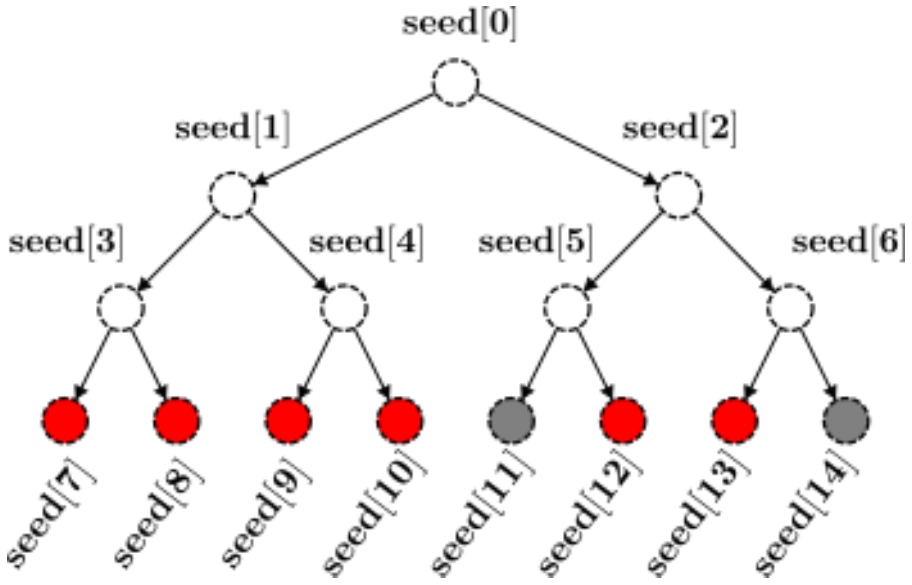


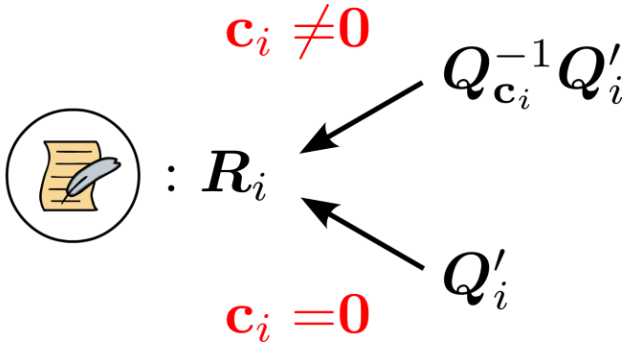
Compression Technique



Suppose **c=(0,0,0,0,1,0,0,2)**

Compression Technique

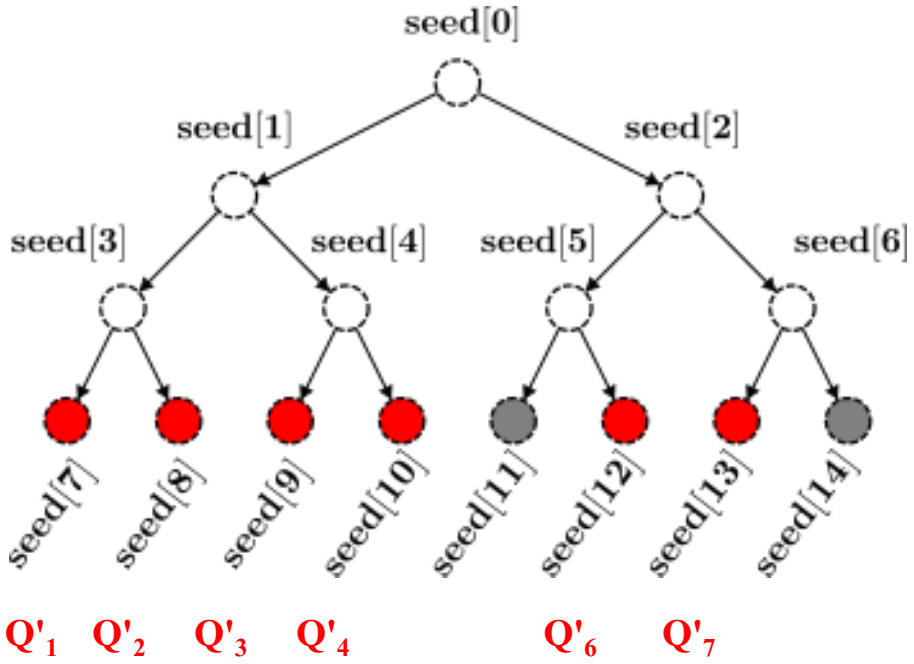


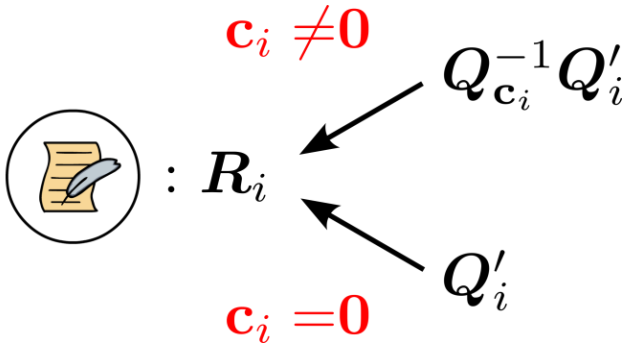


Suppose **c=(0,0,0,0,1,0,0,2)**

● : to publish $\mathbf{Q'_i}$, ● : to hide $\mathbf{Q'_i}$

Compression Technique

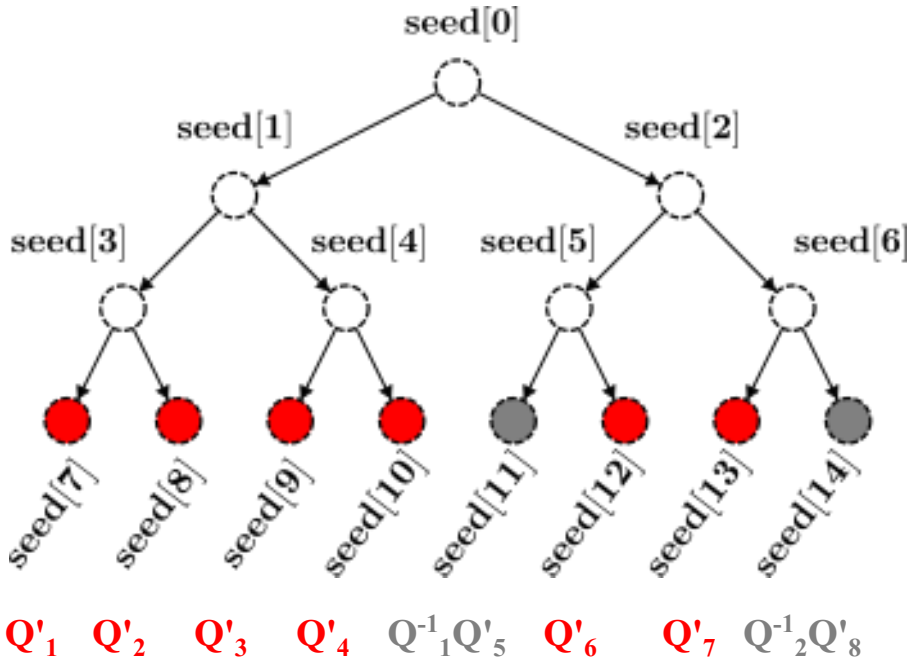


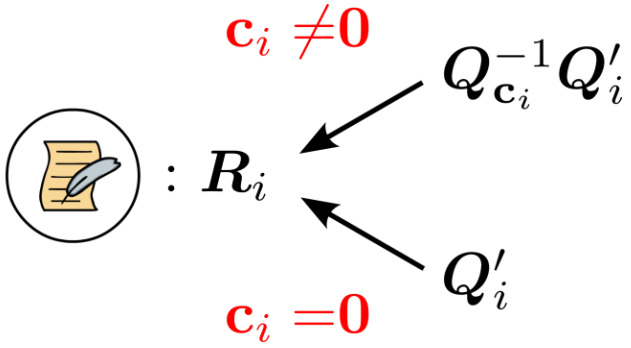


Suppose **c=(0,0,0,0,1,0,0,2)**

●: to publish $\mathbf{Q'_i}$, ●: to hide $\mathbf{Q'_i}$

Compression Technique

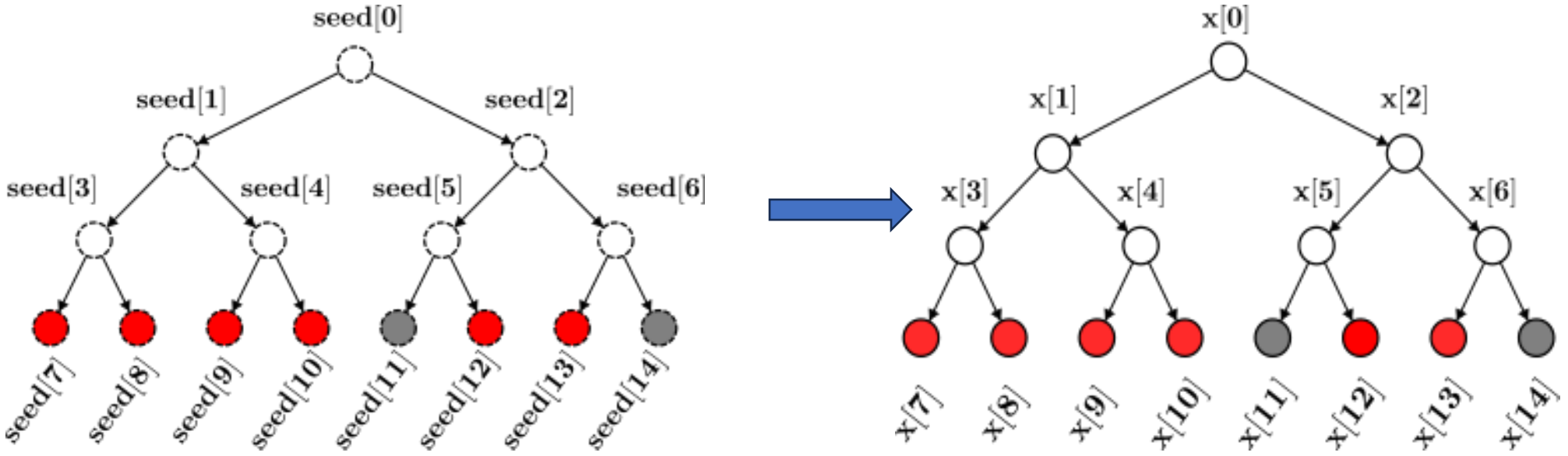




Suppose **c=(0,0,0,0,1,0,0,2)**

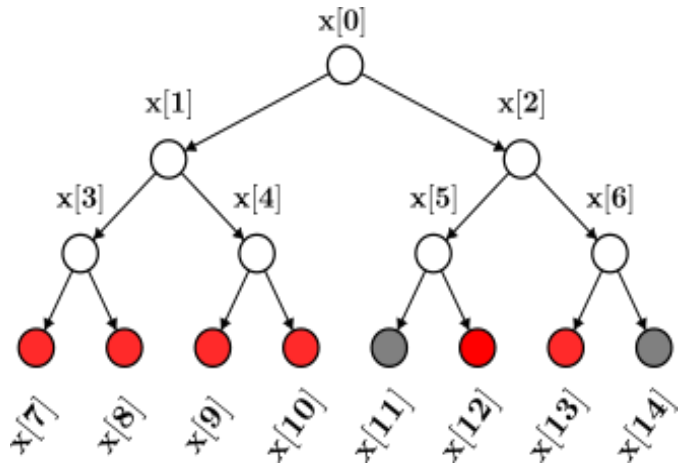
● : to publish **Q'$_i$**, ● : to hide **Q'$_i$**

Compression Technique



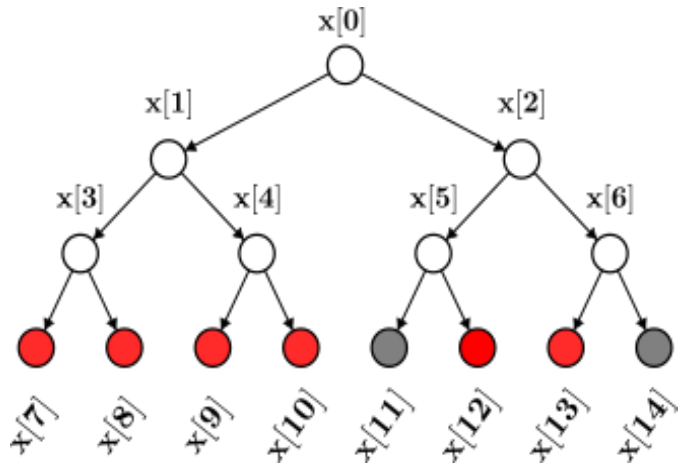
● : to publish $\mathbf{Q'_i}$, ● : to hide $\mathbf{Q'_i}$

Compression Technique



● : to publish $\mathbf{Q'_i}$, ● : to hide $\mathbf{Q'_i}$

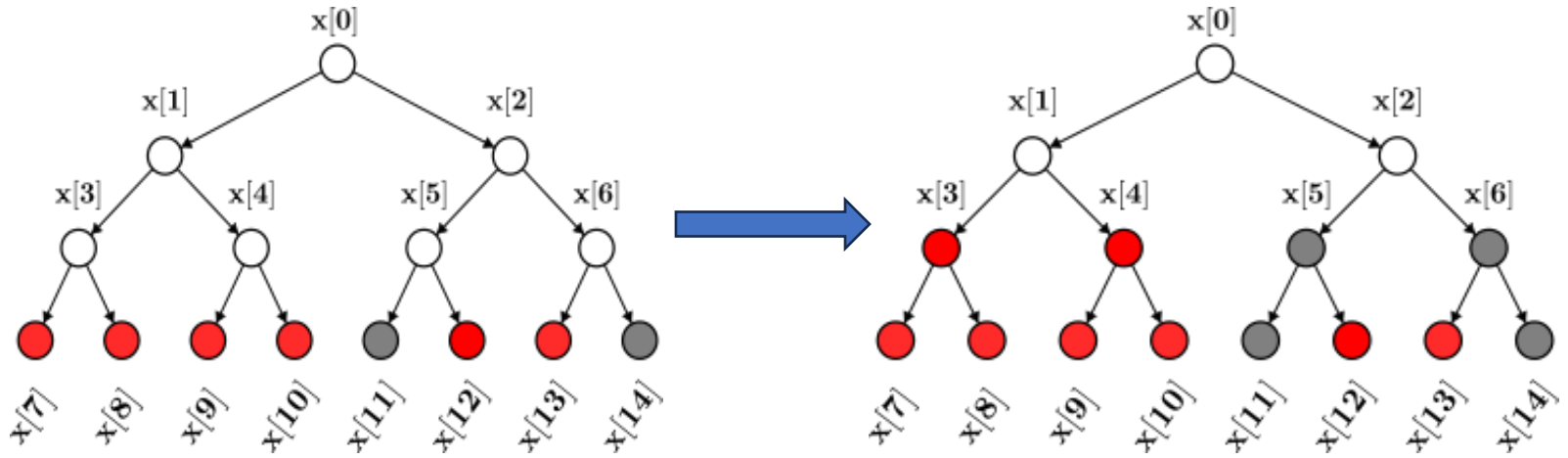
Compression Technique



● : to publish $\mathbf{Q'_i}$, ● : to hide $\mathbf{Q'_i}$

i-th node: ● if both children are ●

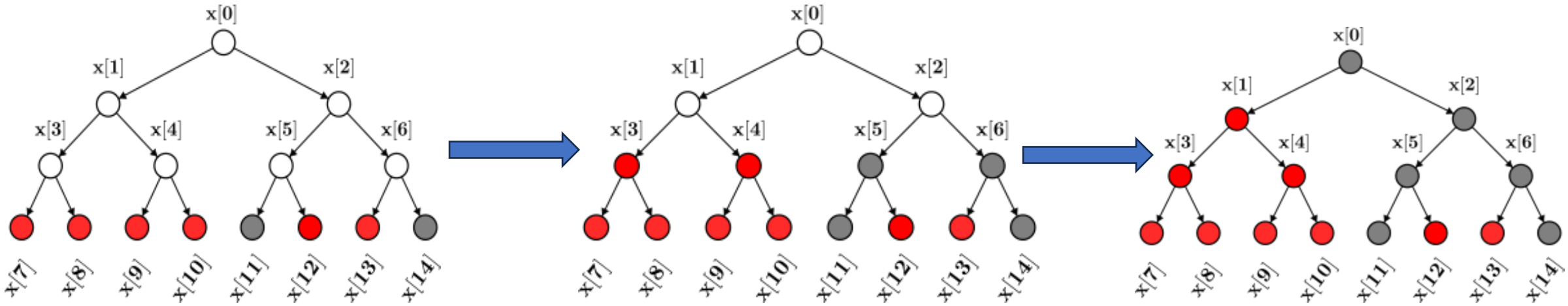
Compression Technique



● : to publish $\mathbf{Q'_i}$, ● : to hide $\mathbf{Q'_i}$

i-th node: ● if both children are ●

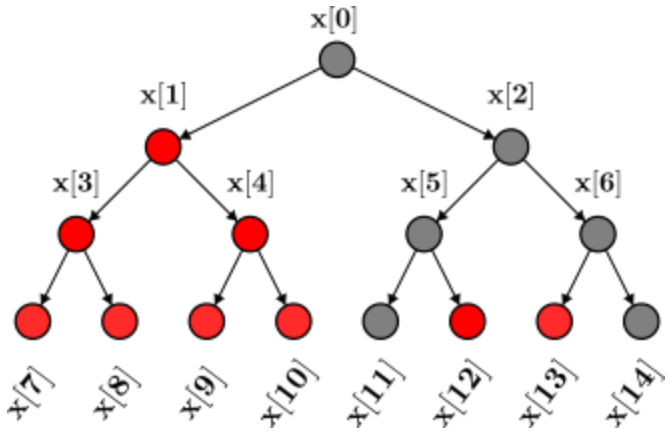
Compression Technique



● : to publish $\mathbf{Q'_i}$, ● : to hide $\mathbf{Q'_i}$

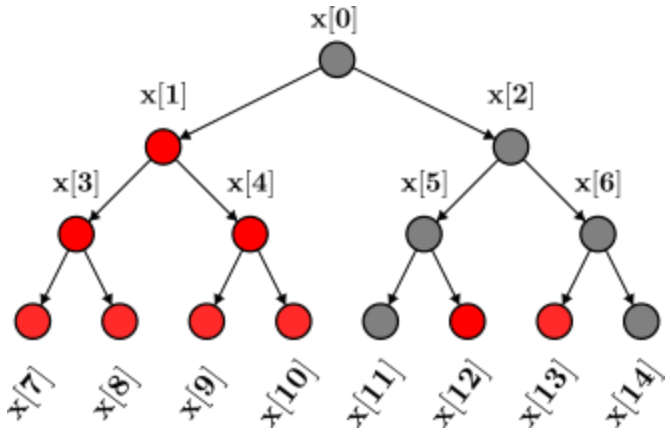
i-th node: ● if both children are ●

Compression Technique



● : to publish $\mathbf{Q'}_i$, ● : to hide $\mathbf{Q'}_i$

Compression Technique

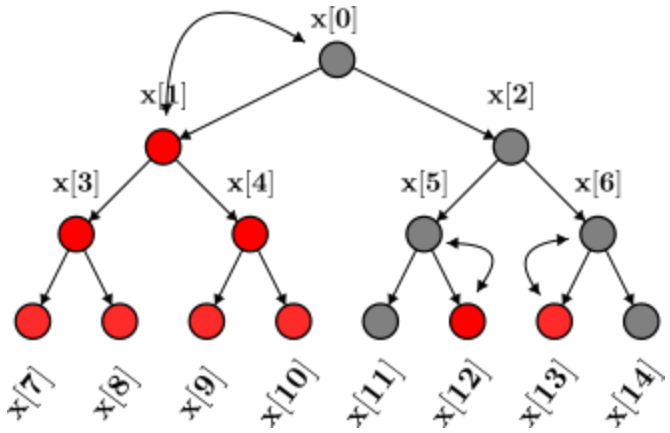


● : to publish $\mathbf{Q'}_i$, ● : to hide $\mathbf{Q'}_i$

Check: i-th node: ● and its parent: ● ?

If yes: publish seed[i]

Compression Technique

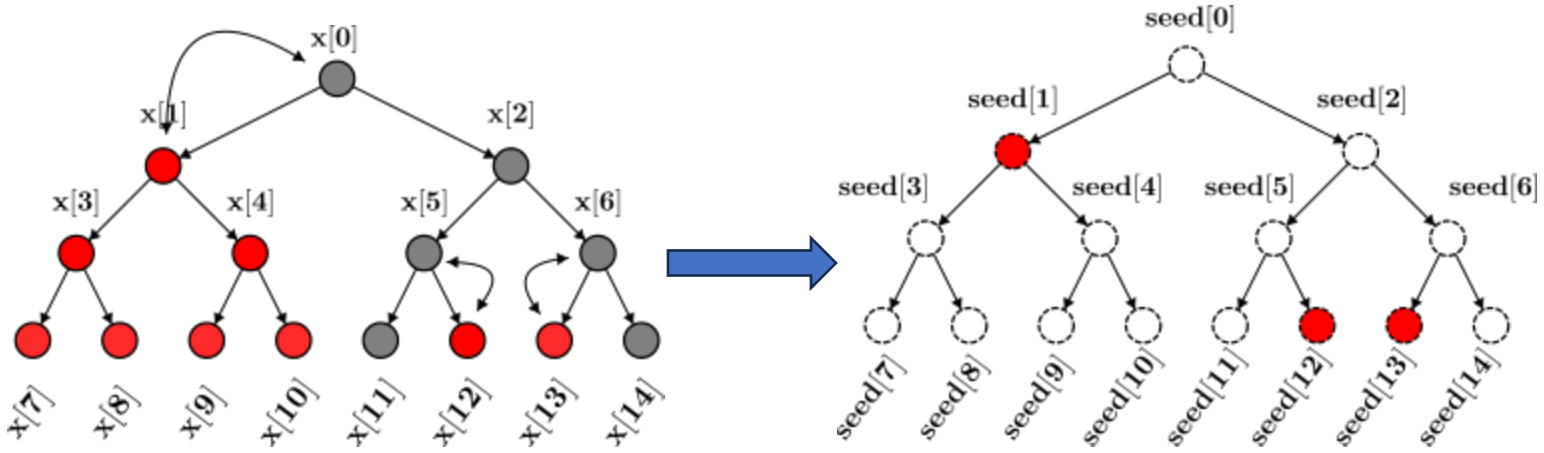


● : to publish $\mathbf{Q'}_i$, ● : to hide $\mathbf{Q'}_i$

Check: i-th node: ● and its parent: ● ?

If yes: publish seed[i]

Compression Technique

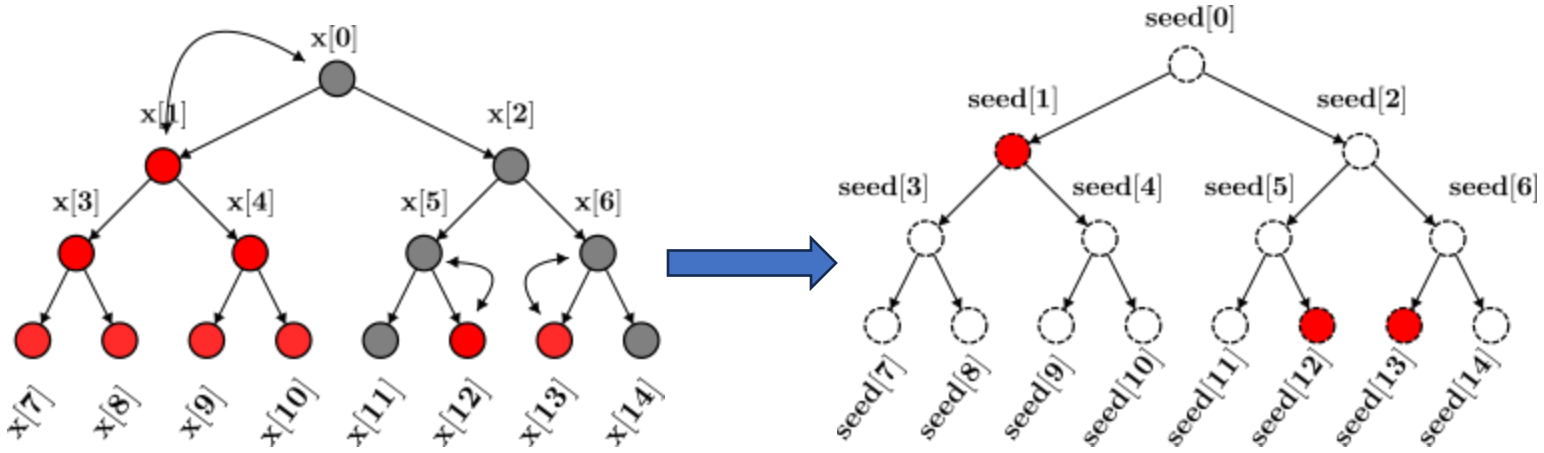


● : to publish $\mathbf{Q'}_i$, ● : to hide $\mathbf{Q'}_i$

Check: i-th node: ● and its parent: ● ?

If yes: publish seed[i]

Compression Technique

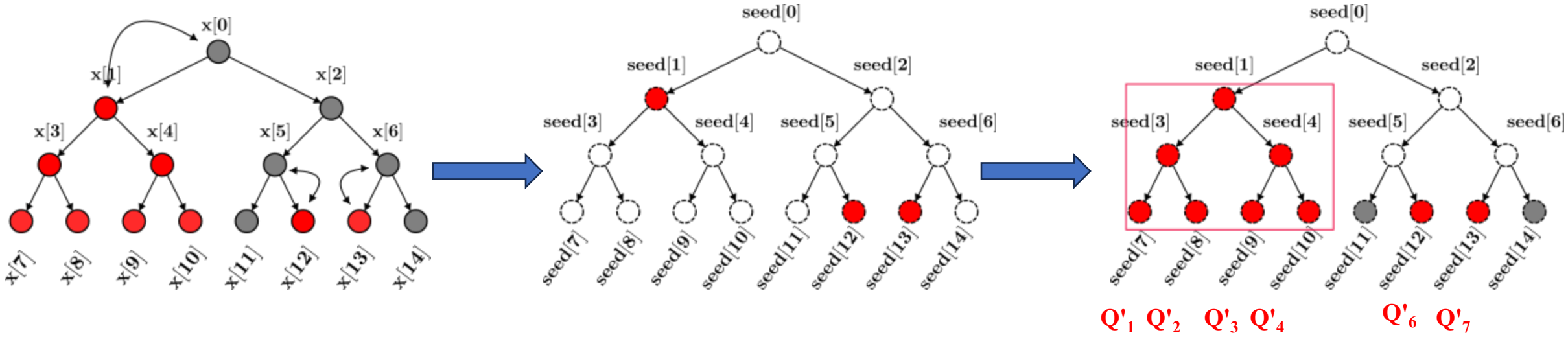


● : to publish $\mathbf{Q'}_i$, ● : to hide $\mathbf{Q'}_i$ publish seed[1],seed[12],seed[13]

Check: i-th node: ● and its parent: ● ?

If yes: publish seed[i]

Compression Technique

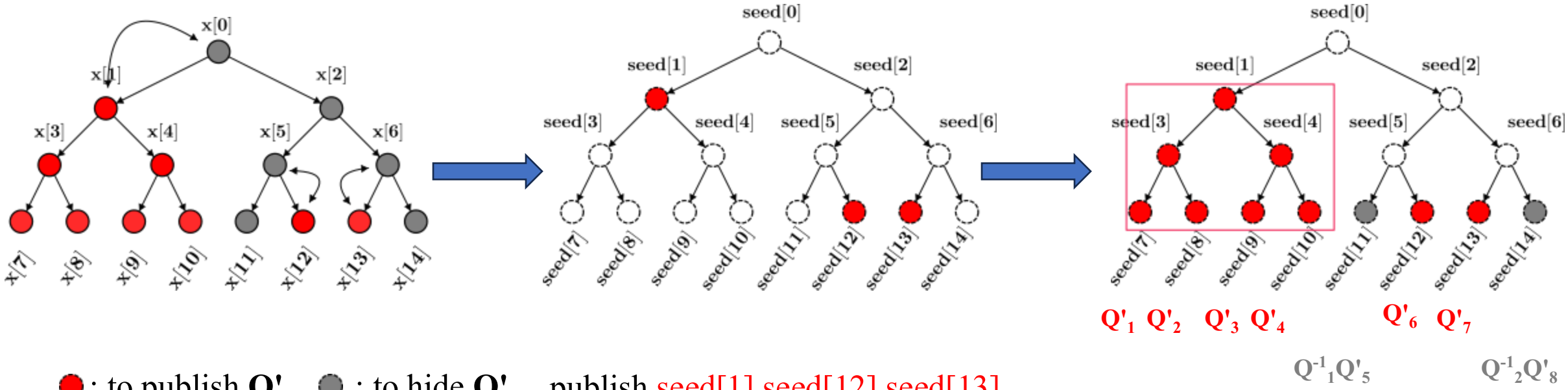


● : to publish $\mathbf{Q'}_i$, ● : to hide $\mathbf{Q'}_i$ publish seed[1],seed[12],seed[13]

Check: i-th node: ● and its parent: ● ?

If yes: publish seed[i]

Compression Technique

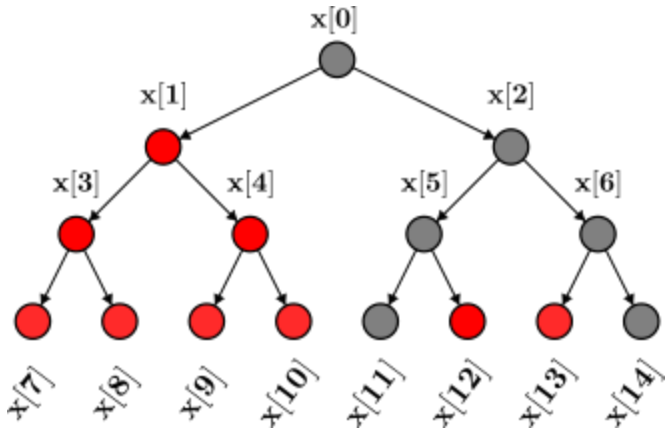


● : to publish $\mathbf{Q'}_i$, ● : to hide $\mathbf{Q'}_i$ publish seed[1],seed[12],seed[13]

Check: i-th node: ● and its parent: ● ?

If yes: publish seed[i]

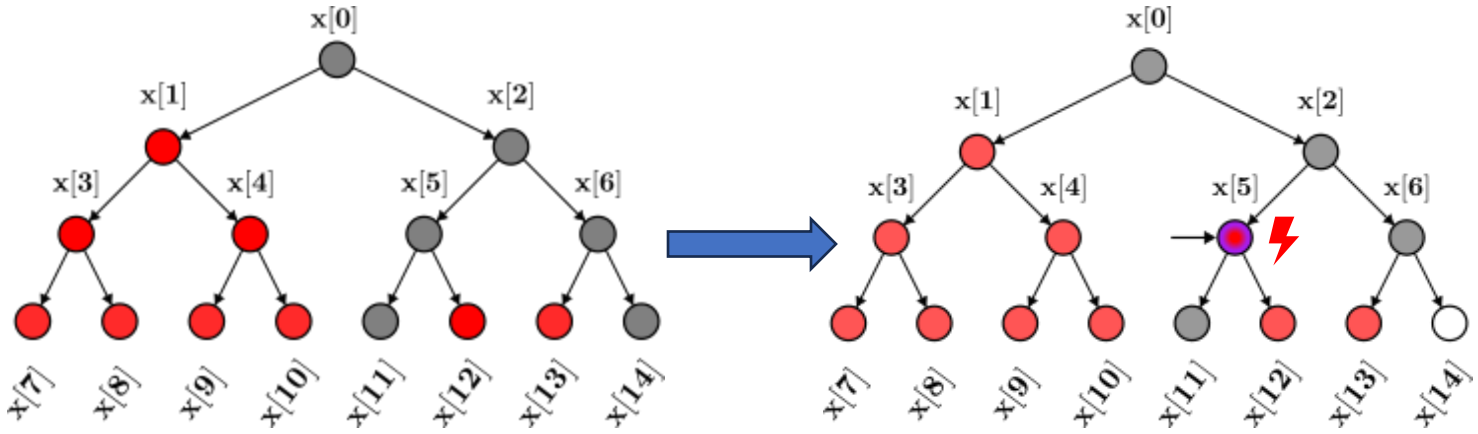
Our Attack Location



Check: i-th node: ● and its parent: ● ?

If yes: publish seed[i]

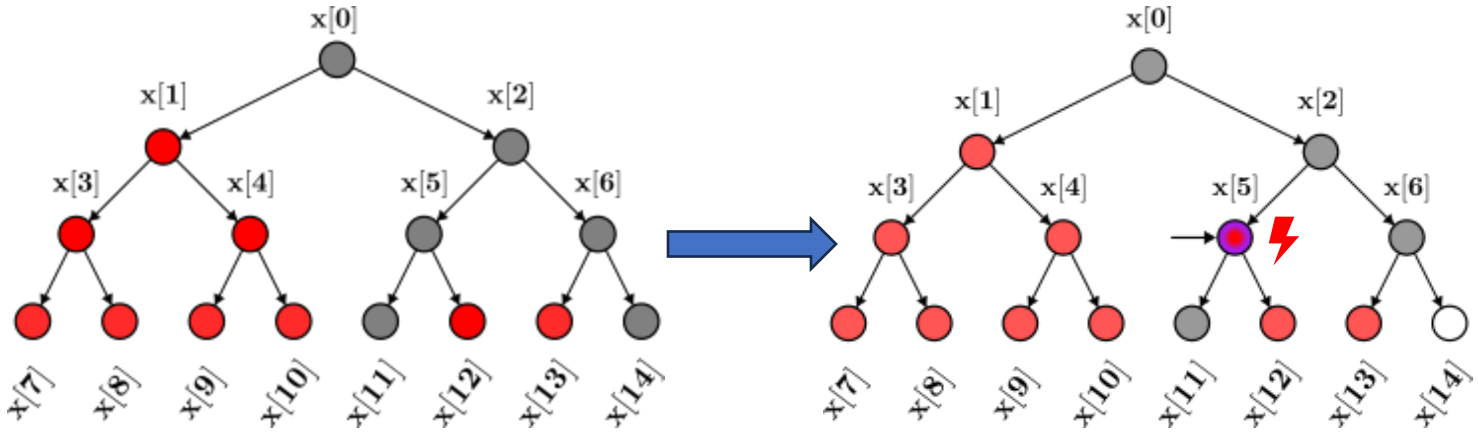
Our Attack Location



Check: i-th node: ● and its parent: ● ?

If yes: publish seed[i]

Our Attack Location

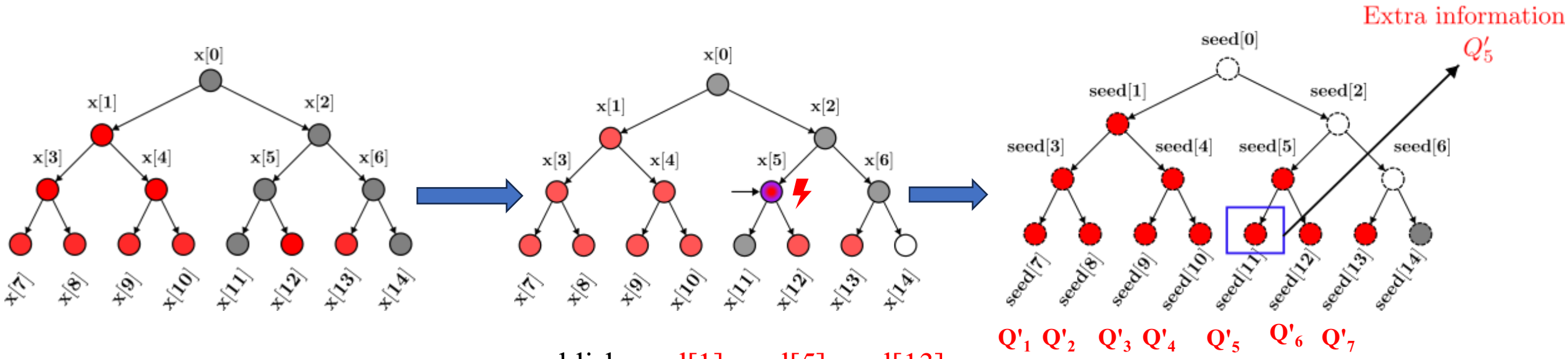


publish seed[1],seed[5],seed[13]

Check: i-th node: ● and its parent: ● ?

If yes: publish seed[i]

Our Attack Location

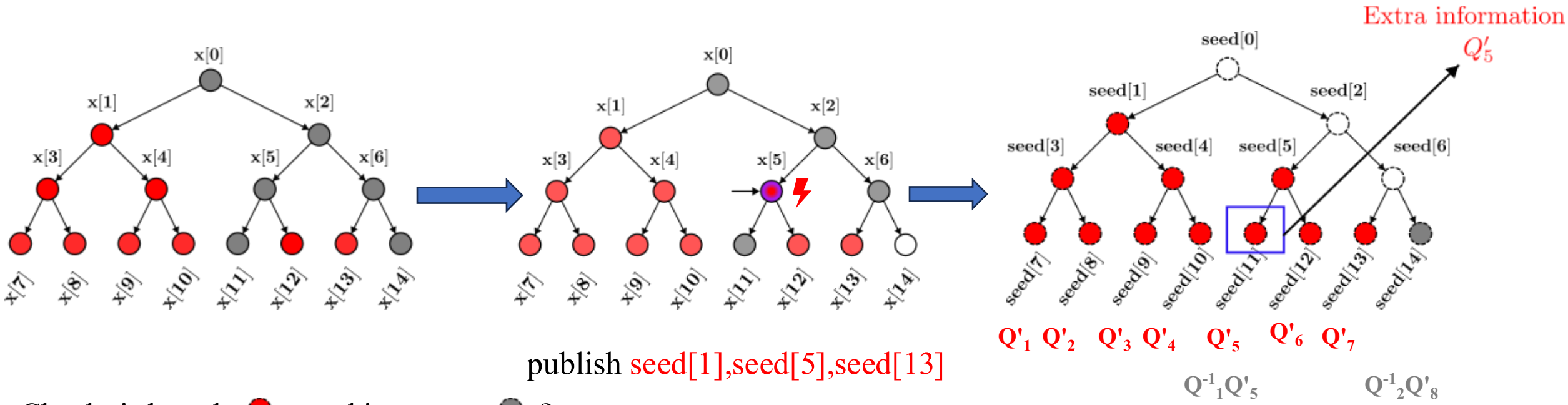


publish seed[1],seed[5],seed[13]

Check: i-th node: ● and its parent: ● ?

If yes: publish seed[i]

Our Attack Location

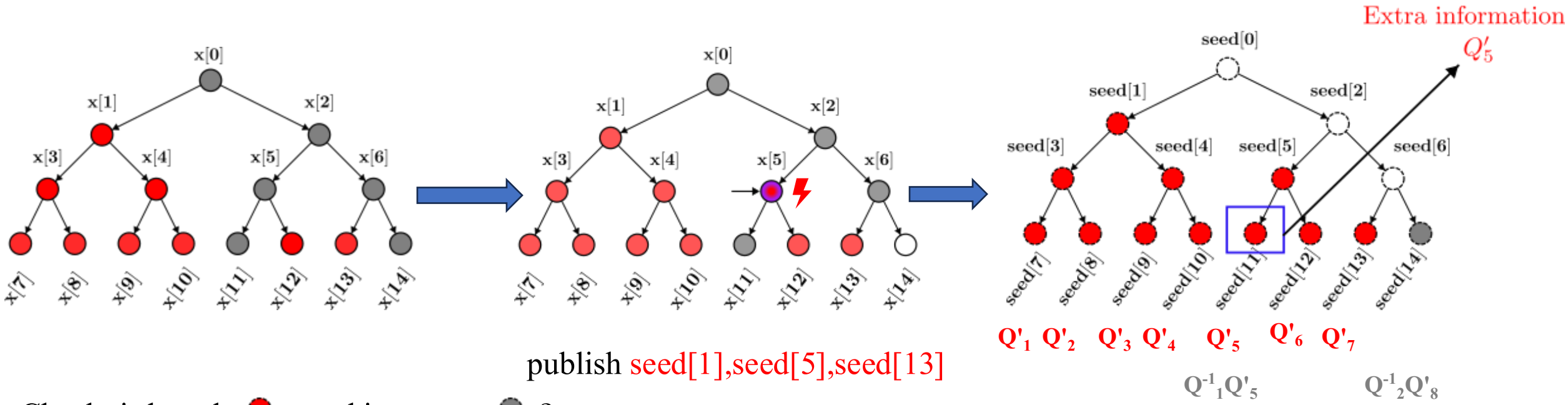


publish seed[1],seed[5],seed[13]

Check: i-th node: ● and its parent: ● ?

If yes: publish seed[i]

Our Attack Location



publish seed[1],seed[5],seed[13]

Check: i-th node: ● and its parent: ● ?

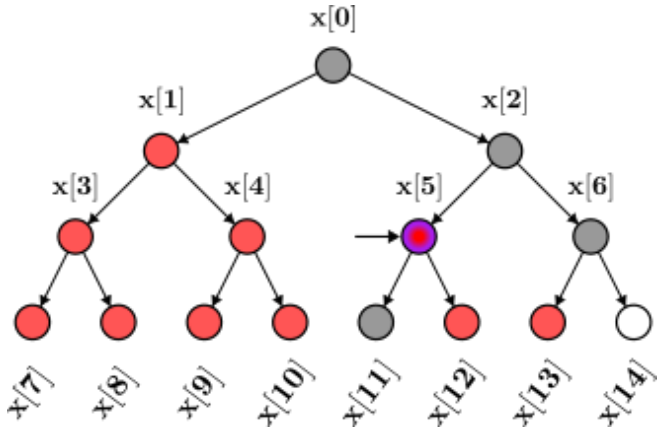
If yes: publish seed[i]

The secret Q_1 can be computed by

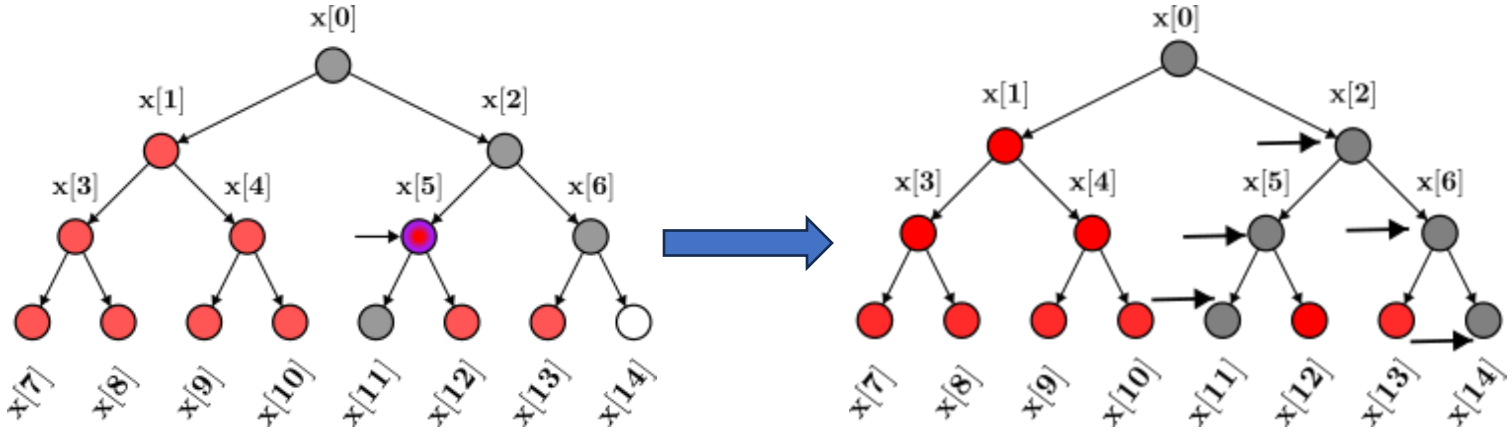
$Q_1=Q'_5(Q^{-1}{}_1Q'_5)^{-1}$

Leaks secret information!!!!

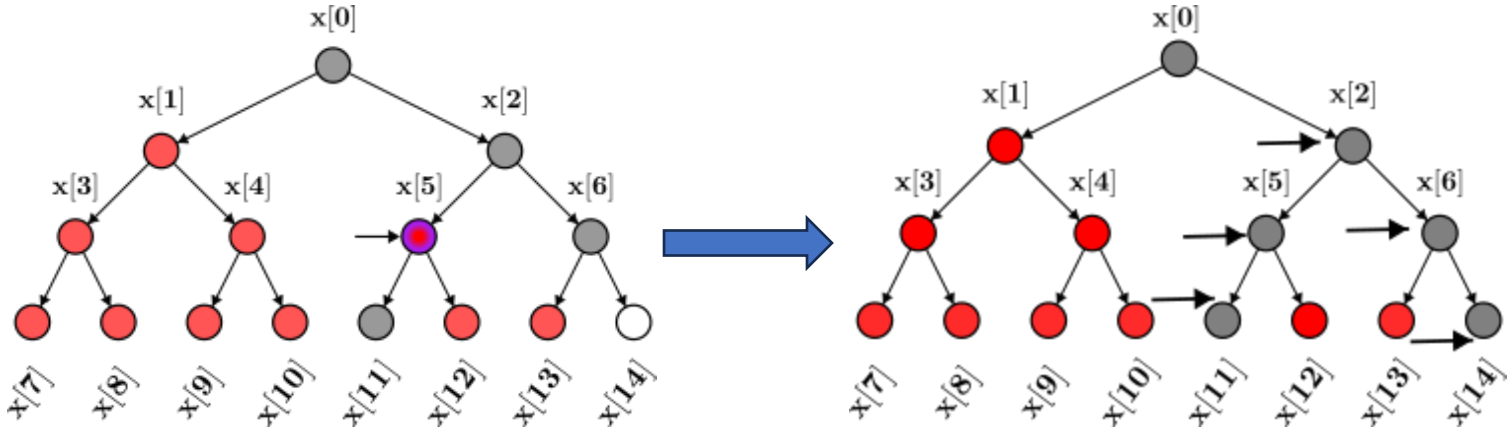
Generic Fault Location



Generic Fault Location

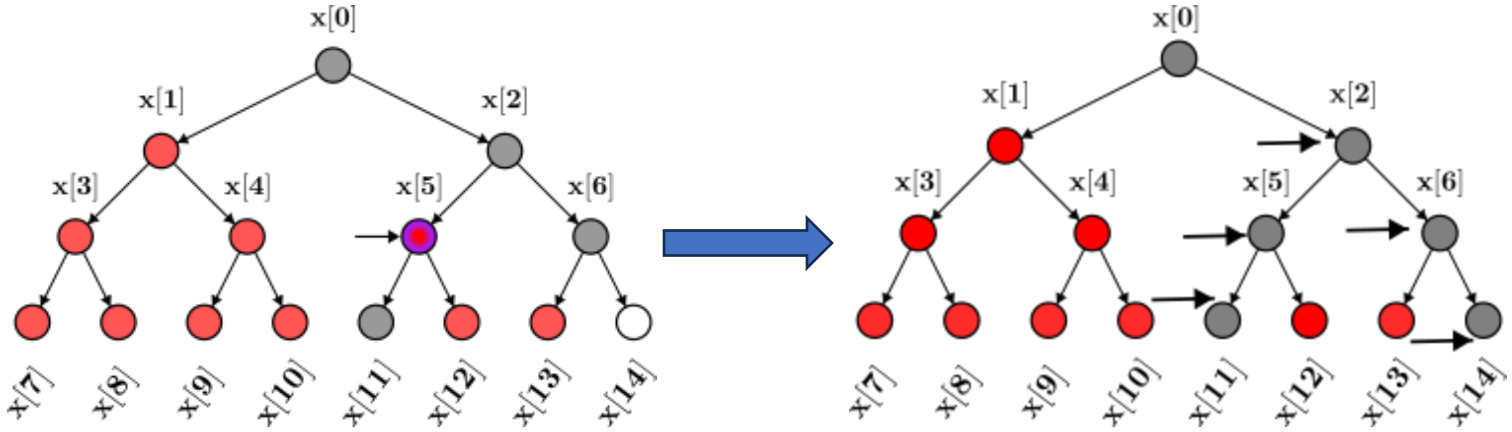


Generic Fault Location



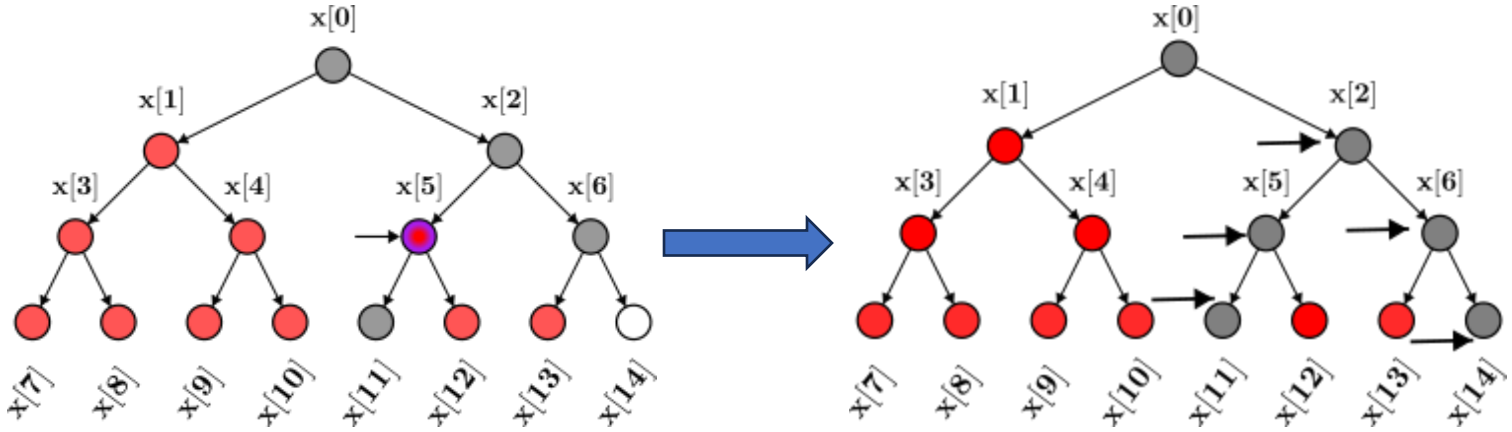
- Change of any grey node by injecting fault will leak the information of the secret

Generic Fault Location



- Change of any grey node by injecting fault will leak the information of the secret
- The nodes with higher heights would always give more informations about secret

Generic Fault Location



- Change of any grey node by injecting fault will leak the information of the secret
- The nodes with higher heights would always give more informations about secret
- For our result, we have targeted the location x[1]/x[2]

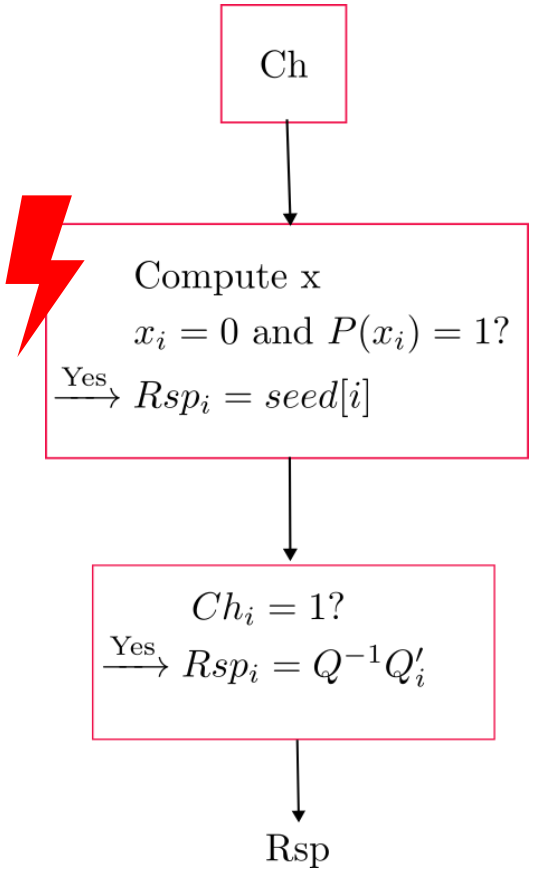
Fault Detection

- After Fault injection we need to answer the questions:
  - Is the fault injected exactly at i-th location?
  - Is the fault injection successful?

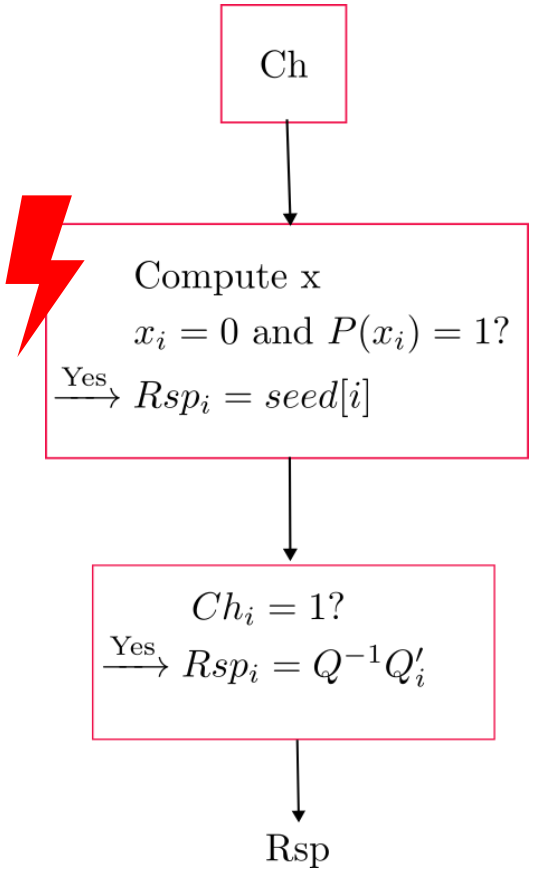
Fault Detection

- After Fault injection we need to answer the questions:
  - Is the fault injected exactly at i-th location?
  - Is the fault injection successful?
- In our work, the fault detection method answers both of the above questions.

How to Prevent the Fault Attack?

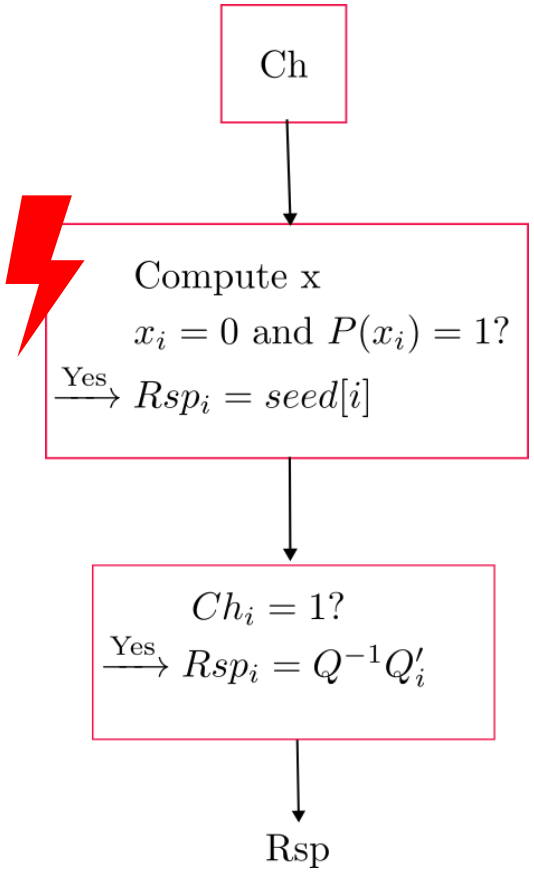


How to Prevent the Fault Attack?



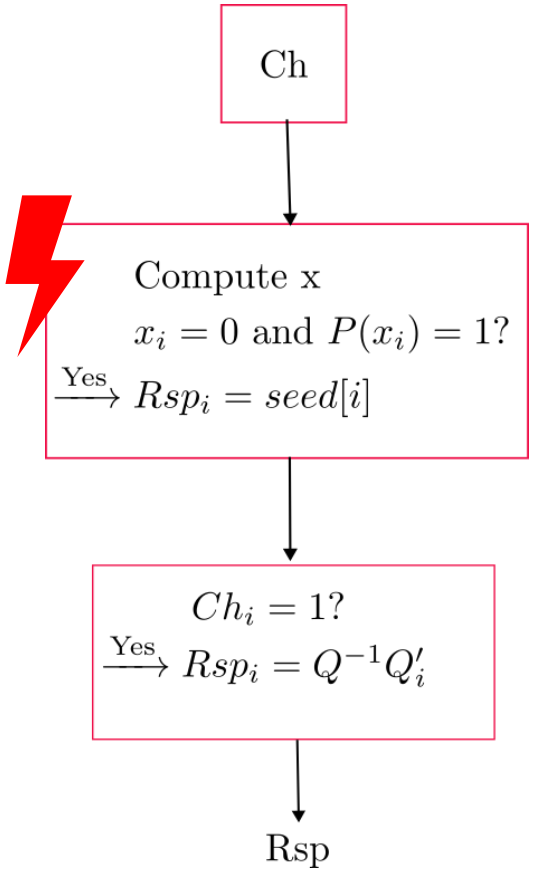
- By injecting fault in the first part to get the value seed[i'] corresponding to Ch_i=1, which will not hamper the second part.

How to Prevent the Fault Attack?



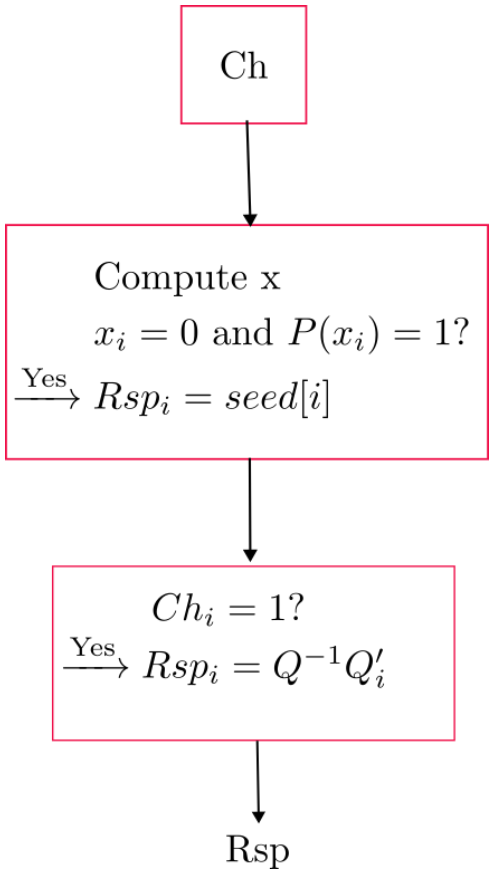
- By injecting fault in the first part to get the value seed[i'] corresponding to Ch_i=1, which will not hamper the second part.
- So, we are getting both values seed[i'] (=Q'_i) and $Q^{-1}Q'_i$ for a non-zero Ch_i.

How to Prevent the Fault Attack?

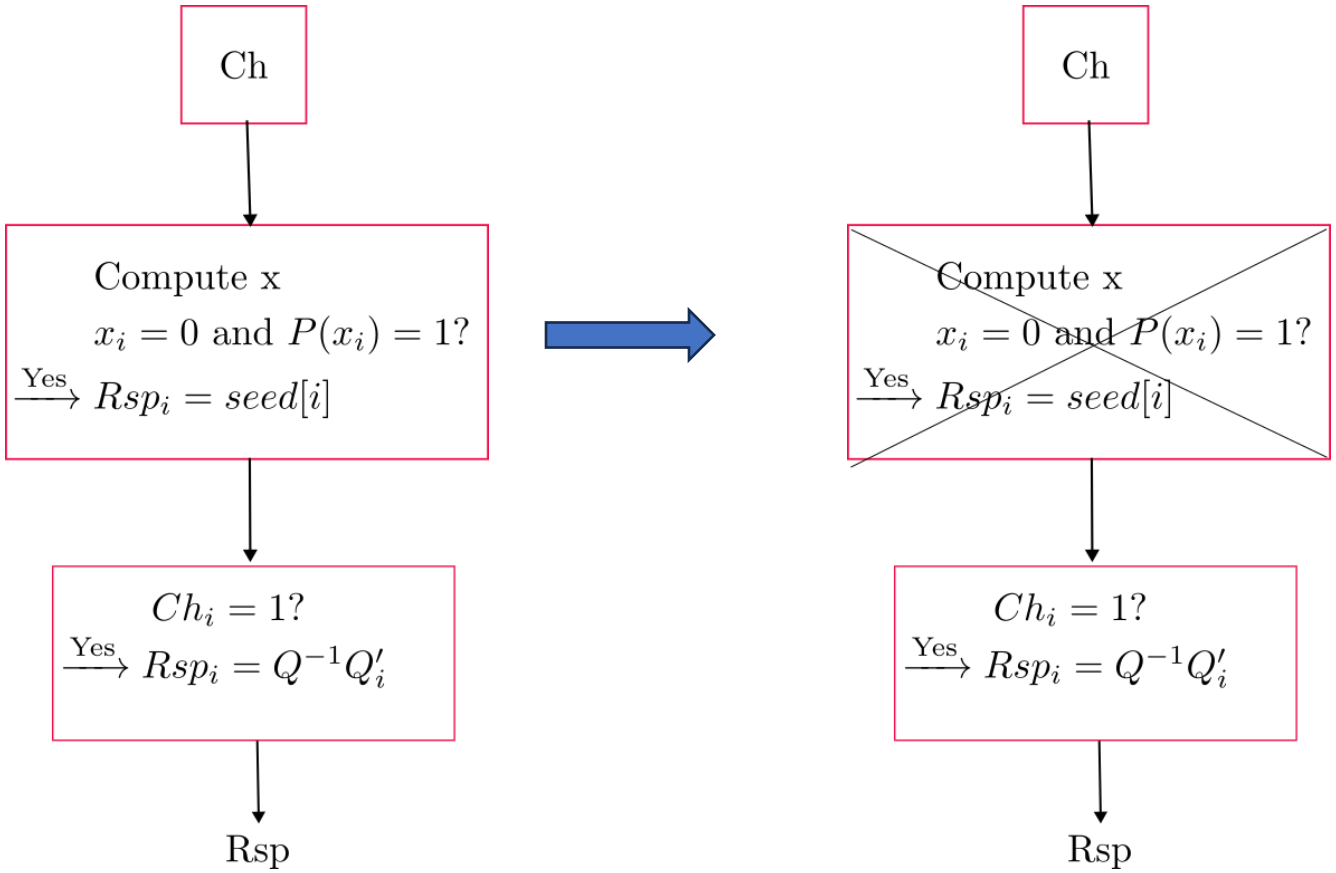


- By injecting fault in the first part to get the value seed[i'] corresponding to Ch_i=1, which will not hamper the second part.
- So, we are getting both values seed[i'] (=Q'_i) and $Q^{-1}Q'_i$ for a non-zero Ch_i.
- The attack successfully done.

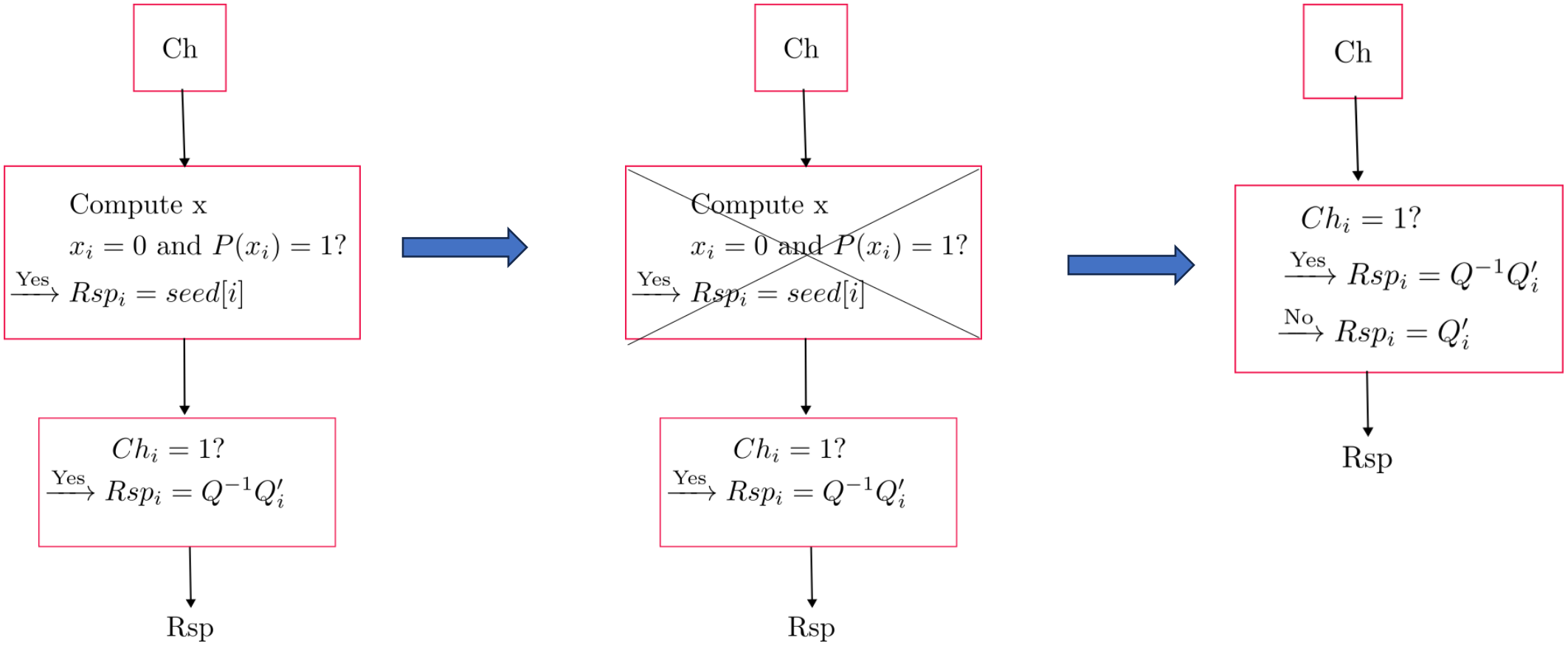
First Countermeasure



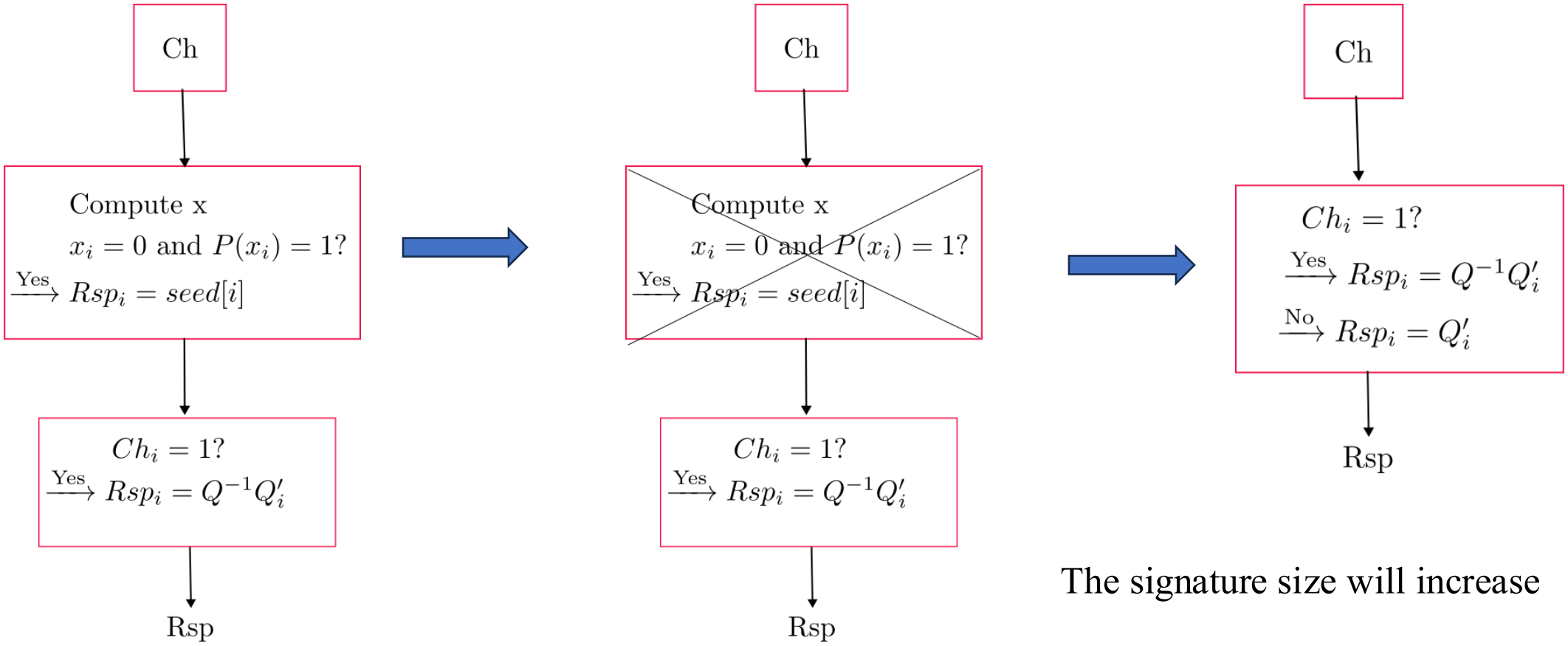
First Countermeasure



First Countermeasure

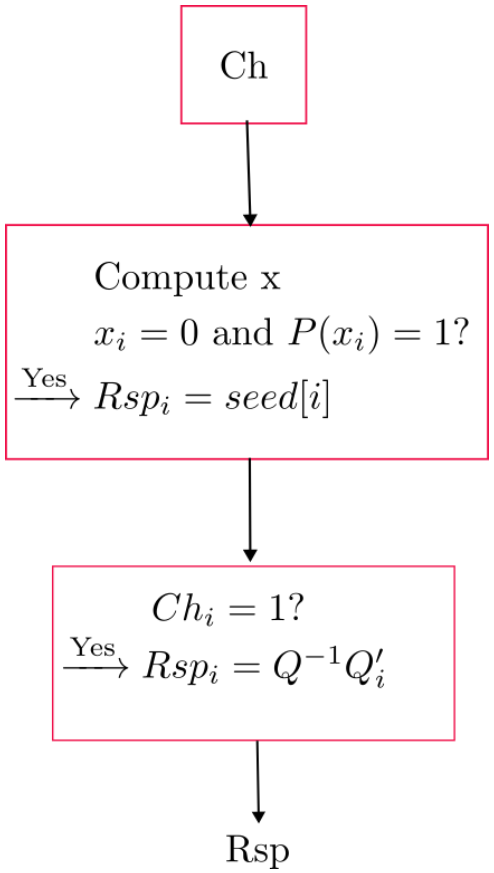


First Countermeasure

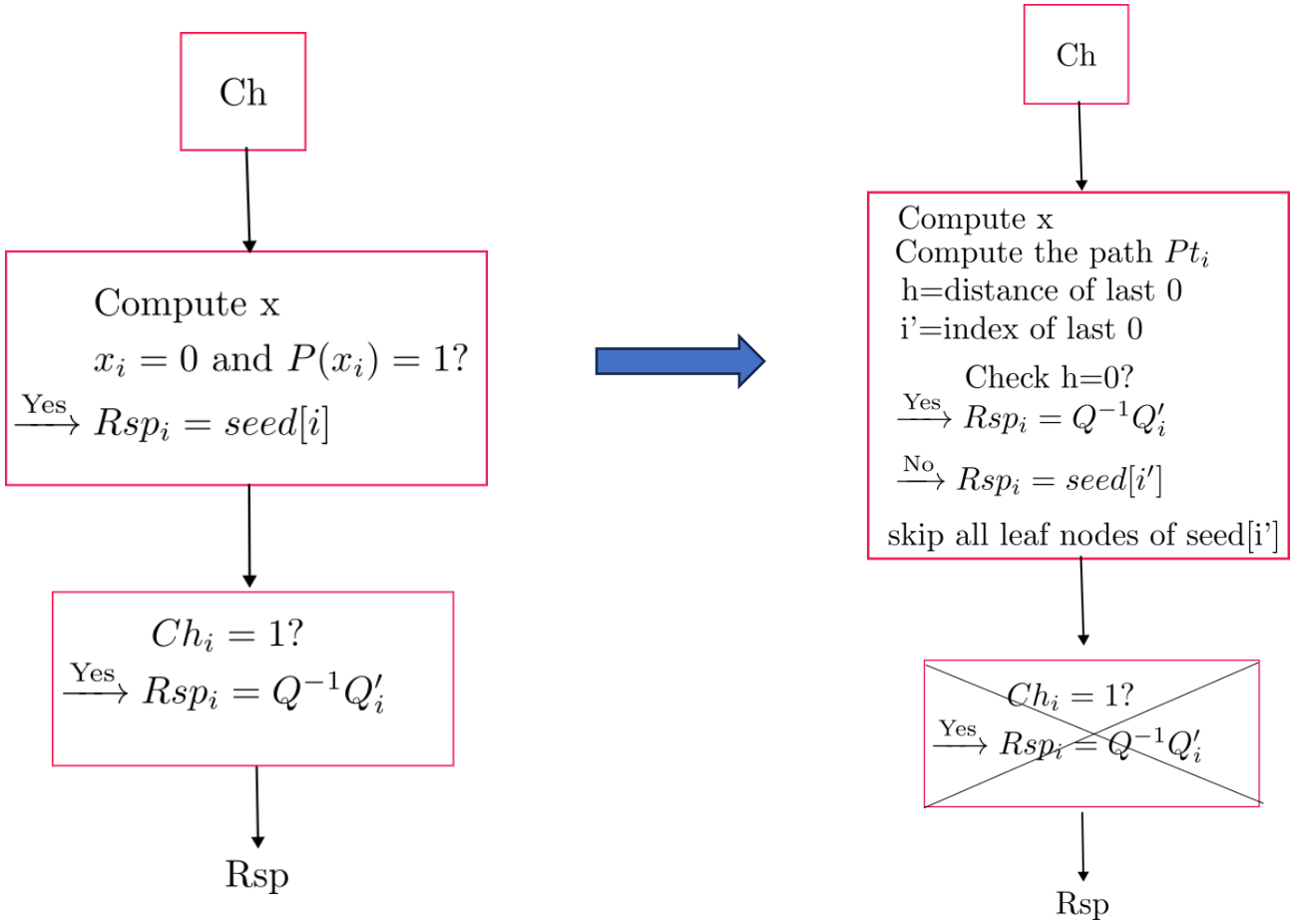


The signature size will increase

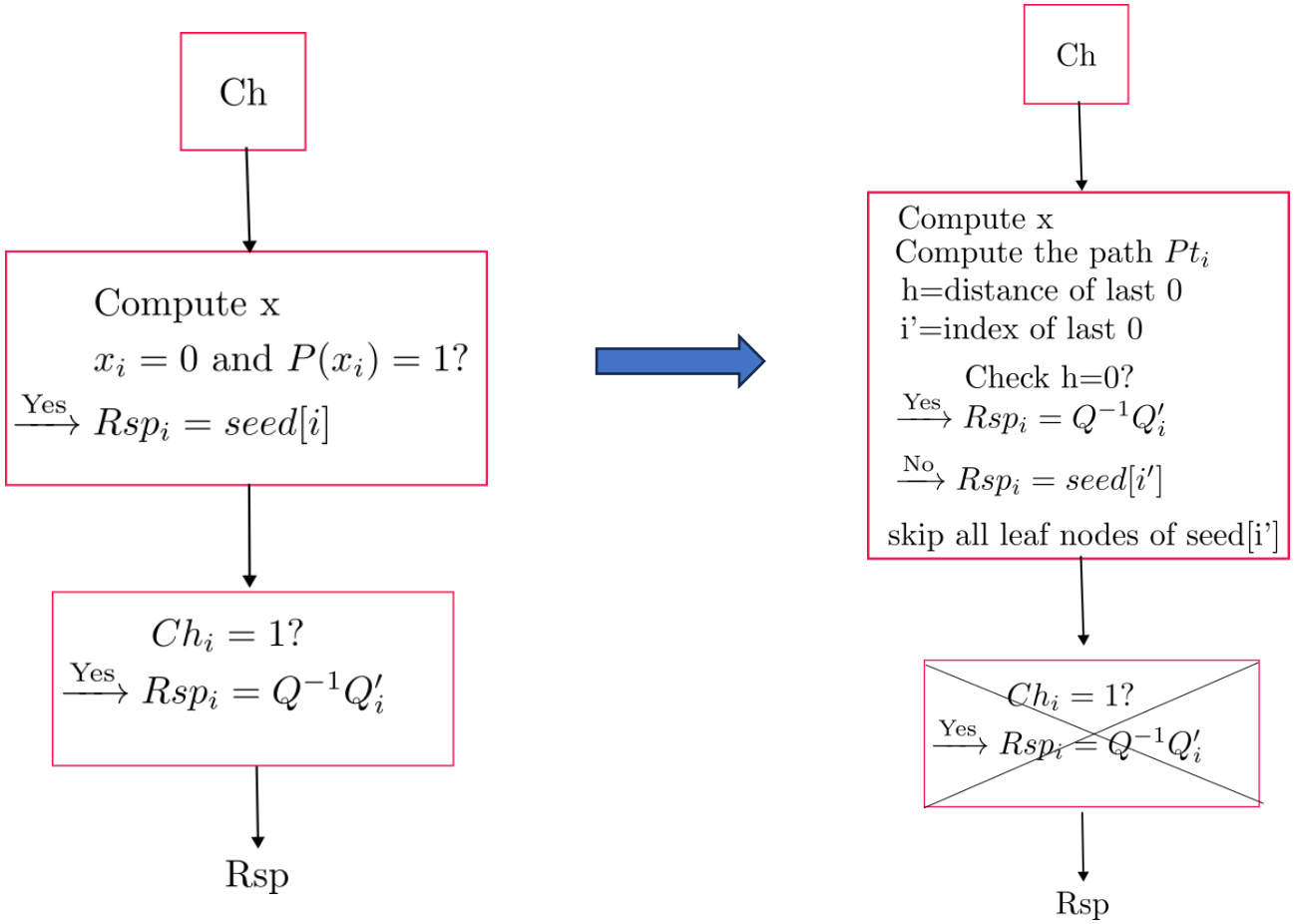
Second Countermeasure



Second Countermeasure



Second Countermeasure



- The signature size remain unchanged

Conclusion

- We show a fault attack on LESS and CROSS.

Conclusion

- We show a fault attack on LESS and CROSS.
- Our proposed fault detection method prevents erroneous secret recovery

Conclusion

- We show a fault attack on LESS and CROSS.
- Our proposed fault detection method prevents erroneous secret recovery
- It can be realized bit flip fault, struck at zero/one fault, instruction skip fault, etc.

Conclusion

- We show a fault attack on LESS and CROSS.
- Our proposed fault detection method prevents erroneous secret recovery
- It can be realized bit flip fault, struck at zero/one fault, instruction skip fault, etc.
- This attack can be applied to other schemes those uses the similar ZK structure
  - ✓E.g. MEDS

Conclusion

- We show a fault attack on LESS and CROSS.
- Our proposed fault detection method prevents erroneous secret recovery
- It can be realized bit flip fault, struck at zero/one fault, instruction skip fault, etc.
- This attack can be applied to other schemes those uses the similar ZK structure
  - ✓E.g. MEDS
- We have proposed two countermeasures that prevent the fault

Questions?

Thank you!